# SciDAC Visualization and Analytics Center for Enabling Technologies
# Mid-stream Program Review


E. Wes Bethel* and Chris Johnson†
*Principal Investigators*

Tom Fogal, Charles Hansen, Valerio Pascucci, Claudio Silva, Allen Sanderson ‡

Sean Ahern, Jeremy Meredith, George Ostrouchov, Dave Pugmire§

Hank Childs, Peer-Timo Bremer, Brad Whitlock¶

Ken Joy, Eduard Deines, Christoph Garth, Bernd Hamann‖

Gunther Weber, Prabhat, Oliver Rübel, Daniela Ushizima, Janet Jacobsen**


April 28, 2009


*Lawrence Berkeley National Laboratory
†Scientific Computing Institute, University of Utah
‡Scientific Computing Institute, University of Utah
§Oak Ridge National Laboratory
¶Lawrence Livermore National Laboratory
‖University of California – Davis
**Lawrence Berkeley National Laboratory

# Contents

## VACET Five-Year Budget Summary

| | FY07 | FY08 | FY09 | FY10 | FY11 | *Total* |
|---|---|---|---|---|---|---|
| LBNL | 344325 | 344325 | 344325 | 344325 | 344325 | 1721625 |
| LLNL | 656,817 | 656,817 | 328817 | 328817 | 328817 | 2300085 |
| ORNL | 374221 | 374221 | 374221 | 374221 | 374221 | 1871105 |
| UCD | 298991 | 298991 | 298991 | 298991 | 298991 | 1494955 |
| Utah | 525,646 | 525,646 | 853646 | 853646 | 853646 | 3612230 |
| *Total* | 2200000 | 2200000 | 2200000 | 2200000 | 2200000 | 11000000 |

## VACET Five-Year Staffing Summary

| | FY07 | FY08 | FY09 | FY10 | FY11 |
|---|---|---|---|---|---|
| LBNL | Bethel, Weber, Prabhat, Aragon, Ruebel (S), Gosink (S) | Bethel, Weber, Prabhat, Aragon, Ushizima (PD), Ruebel (S), Gosink (S) | Bethel, Weber, Prabhat, Ushizima (PD), Ruebel (S), Gosink (S) | Bethel, Weber, Prabhat | Bethel, Weber, Prabhat |
| Utah | Johnson, Hansen, Silva, Parker, Sanderson | Johnson, Hansen, Silva, Parker, Sanderson, Cole | Johnson, Hansen, Silva, Pascucci, Sanderson, Fogal | Johnson, Hansen, Silva, Pascucci, Sanderson, Fogal | Johnson, Hansen, Silva, Pascucci, Sanderson, Fogal |
| LLNL | Pascucci, Childs, Whitlock, Bremer | Pascucci, Childs, Whitlock, Bremer | Childs, Whitlock, Bremer | Childs, Whitlock, Bremer | Childs, Whitlock, Bremer |
| ORNL | Ahern, Meredith, Ostrochov | Ahern, Meredith, Ostrochov, Pugmire | Ahern, Meredith, Ostrochov, Pugmire | Ahern, Meredith, Ostrochov, Pugmire | Ahern, Meredith, Ostrochov, Pugmire |
| UCD | Joy, Hamman | Joy, Hamman, Garth (PD), Budge (S), Anderson (S), Krishnan (S) | Joy, Hamman, Garth (PD), Deines (PD), Budge (S), Camp (S), Krishnan (S) | Joy, Hamman, Garth (PD), Deines (PD), Agranovsky (S), Krishnan (S) | Joy, Hamman, Garth (PD), Deines (PD) |

(S) = Student, (PD) = Postdoc

Likely missing: Students from Utah, Davis

# 1 Executive Summary

## 1.1 SciDAC Program Goals

Initiated in 2001 and continued in 2006, the SciDAC program objectives are aimed at enabling new science through the use of today's largest computational platforms. The SciDAC program portfolio spans: (1) science application teams, which aim to produce new scientific results in specific science domains; (2) centers for enabling technologies, which provide production-quality, parallel capable software infrastructure aimed at enabling science at the petascale; (3) institutes, which have a long-term research focus and include an outreach and training element; (4) partnerships, which are teams of domain, computer, and mathematics scientists focused on solving a specific science application problem.

As a whole, the SciDAC portfolio supports a diverse spectrum of activities that target subsets of the overall mission, "enable science on petascale platforms." At the crossroads of the many science domains are mathematics and computer science, which provide the underpinnings that in turn are adapted for use in specific science projects. Per the SciDAC2 call for proposals, "the key to the success of the SciDAC program has been the power of multidisciplinary teams that bring together experts in the scientific discipline, computer science and applied mathematics."

The overriding SciDAC program objective, "enable science at the petascale," offers many opportunities and research challenges. One is to make progress towards realizing peak sustained computational rates on petascale platforms. Another, and one that is more the focus of this document, is the fact that these larger codes produce ever larger and complex results: an urgent challenge that motivates our work is enabling scientific knowledge discovery through visual data analysis.

## 1.2 VACET Mission

The SciDAC Visualization and Analytics Center for Enabling Technologies (VACET) focuses on leveraging scientific visualization and analytics software technology as an enabling technology for increasing scientific productivity and insight. Our mission is to foster scientific insight through creating and deploying effective data understanding technology that is truly responsive to the needs of our stakeholders in the scientific research community who are "awash in data." It is widely accepted that one of the bottlenecks in contemporary science is the need to gain insight from vast collections of complex data.

The vision for our Center is to respond directly to this challenge by adapting, extending, creating when necessary and deploying visualization and data understanding technologies for our science stakeholders. Organized as a Center for Enabling Technologies, we are well positioned to be responsive to the needs of a diverse set of scientific stakeholders in a coordinated fashion using a range of visualization, mathematics, statistics, computer and computational science and data management technologies.

## 1.3 VACET Accomplishments Meet SciDAC Program Objectives

As we reach the halfway point in our five-year project, we are pleased to report that VACET is realizing exemplary success in all program focus areas. These accomplishments help the SciDAC program achieve its primary mission, which is to enable scientific discovery through the use of petascale class computing platforms. Our accomplishments span science impact (Section 3), production-quality petascale capable visual data analysis software infrastructure that is being widely adopted by the science community (Section 3) and is helping the SciDAC Program to show

how science is enabled at the petascale (Section 5), field-leading and award-winning visualization research (Sections 4 and 9), positive impact on the field of visualization (Section 4), effective use of ASCR computing resources (Section 10), outreach to and effective parterning with the science community (Section 2), and training future generations of visualization scientists through our internal organization and operation model (Section 7). Our team organization and ongoing operation is highly successful and efficient; it is an example of a well run, efficient and productive SciDAC Center (Section 7).

#### 1.3.1 Positive Science Impact

From the outset, our primary mission and objective has been to enable new science. We have made excellent progress on this broad mission objective along several different fronts.

First, our team has provided new capabilities to science teams that aid in knowledge discovery. They are able to see, for the first time, features and attributes that were formerly hidden from view. An excellent example is our topological analysis work (Section 4.4) that resulted in our science collaborators gaining new insight into the fundamental processes of combustion. Such discoveries and understanding are simply not possible by "making a picture of the simulation data." Instead, these discoveries are enabled by fundamental advances in analysis technology.

Other examples include: (1) the new capability for accelerator researchers to see for the first time all particles that meet a minimum level of being "scientifically interesting" in 3D and in conjunction with multi-modal visual presentation (Section 3.1); (2) multi-modal (traditional CFD variables, vector-valued magnetic field; multi-frequency radiation transport) visual data exploration of supernova simulation results from a petascale class machine (Section 3.3); (3) ultra-high resolution climate models run on petascale class platforms (Sections 3.5 and 3.6) contain new features not visible since these massive datasets are not typically processable using "standard" desktop visualization applications; (4) as simulation codes evolve, they often adopt new forms of computational grids to achieve higher levels of efficiency on parallel platforms (e.g., mapped-grid AMR in fusion science (Sections 3.12 and 8) and geodesic grids in climate science (Section 3.5)) – our team has provided the new capability for science teams to perform visual data analysis and exploration on these unusual computational grids.

A second measure of science impact is cost savings. Through a collaboration with VACET, the SciDAC Applied Partial Differential Equations Center (APDEC) realizes a direct cost savings to their project: formerly, they built and maintained their own in-house AMR visualization application since there was no alternative. Our interactions with APDEC focused on first identifying their AMR visualization requirements then adapting and extending VisIt to support those needs (Section 3.12), allowing them to make the transition to a "buy rather than build" product that meets their needs and that will allow them to perform visual data analysis on significantly larger datasets in the future due to VisIt's parallel processing capabilities.

> Prior to working with VACET, our team developed and maintained an in-house adaptive mesh refinement (AMR) visual data analysis software application (ChomboVis). VACET helped us to transition from ChomboVis to VisIt, a production-quality, parallel capable visual data analysis application. . . . In addition to a direct cost savings from not having to support and maintain such software, our team is in a better position now to perform visual data analysis and knowledge discover of complex, time-varying AMR datasets now and in the future as we move into the petascale regime of computing and data production. Having this kind of capability is crucial for the success of our project and for our science stakeholders, who have either adopted or are transitioning to day-to-day use of VisIt. – *Phillip Colella, LBNL*

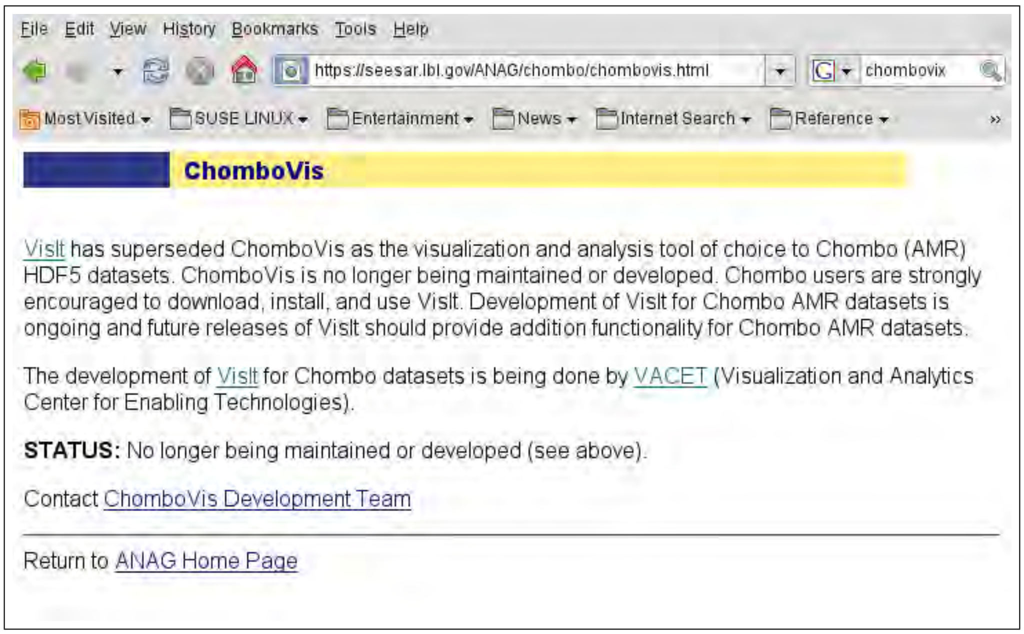



Figure 1: The web page for the SciDAC Applied Partial Differential Equations Center (APDEC) in-house ChomboVis visualization application indicates, as of approximately September 2008, that VACET's VisIt application is the tool of choice for use with Chombo-produced AMR datasets.

The APDEC story is a "poster child" for VACET's success as a Center for Enabling Technology in providing production-quality, parallel capable visual data analysis software infrastructure. There are numerous other examples of science teams headed in this direction: computational astrophysics (Sections 3.2, 3.3), climate modeling (Sections 3.5, 3.6), fusion (Sections 3.9, 3.10, and 3.11). This theme is present in many of the support letters included with this report (Section 11).

A third measure of science impact is improving efficiency reducing the duty cycle in the scientific inquiry process. One of our projects resulted in reducing the duty cycle of a scientific inquiry from hours to seconds (Section 3.1), thereby increasing scientific efficiency. That particular accomplishment was made possible through a close collaboration withe SciDAC Scientific Data Management Center. Furthermore, the software infrastructure we developed to achieve this breakthrough, which includes advances in user interface technology, high performance parallel data I/O, and use of the SDM Center's FastBit software, is slated to be included in the next public release of VisIt so as to be accessible by the worldwide science community.

More broadly, improved efficiency can result from increasing use of parallelism on the part of visual data analysis software infrastructure. To that end, our team has ported VisIt to run on `franklin` at NERSC/LBNL and `jaguar` at LCF/ORNL (see testimonials from the directors of DOE's open computing facilities in Section 11). Through outreach activities, including tutorials directly with science project teams (Section 9.3) and participation in numerous workshops (Section 9.6), we are helping individual science teams and the broader scientific community gain first-hand understanding of how to use leading visualization technology to solve their challenging scientific data understanding problems.

Effective use of parallel resources requires effort above and beyond simply parallelizing the algorithms and execution model of the visualization application. It also requires design and engineering of effective parallel I/O, which can often be one of the most resource-consuming stages of the visualization pipeline. Our team's efforts in this area include: (1) direct support to a SciDAC climate science team to design and implement data model and parallel I/O library for use on the Cray XT4 platforms at NERSC/LBNL and LFC/ORNL (Section 3.5); (2) direct support to a SciDAC fusion science center to enable parallel I/O of fusion (and accelerator) simulation data (Section 3.9) where that effort will "spill over" to help numerous other projects that now use or will use those same simulation codes; (3) coordination with support staff at the SC centers to optimize and tune I/O on petascale platforms (Section 3.5). It turns out that many simulation projects initially focus on achieving good parallelization and scalability, and it isn't until later in the project that they realize they must also address parallel I/O, data models, and file formats.

A related view of efficiency can be couched in terms of providing new capabilities within familiar toolsets. VACET has collaborated with the Earth System Grid project to deliver important new 3D visual data analysis capabilities that are included with the familiar ESG toolset named the Climate Data Analysis Toolkit (CDAT). Our contributions are fully integrated with the CDAT framework and now in the hands of a large body of climate scientists (Section 3.4). Additionally, VACET has "productized" an award-winning technology[1] that is also included with CDAT. This new technology helps simplify and streamline complex visual data exploration and analysis tasks (see Section 8).

#### 1.3.2 Production Quality, Petascale Capable Software Infrastructure

One of the main objectives for SciDAC Centers is to deliver production quality, parallel capable software infrastructure that enables science at the petascale. This objective is one of the central mission thrusts for VACET, and to that end, we have exceeded our initial expectations. Our production quality, parallel capable visual data analysis software application runs on all modern parallel platforms, including the Cray XT4/XT5 systems at NERSC/LBNL and LCF/ORNL, as well as all auxiliary systems at DOE's open computing facilities that provide post-processing analysis capabilities. Our software not only runs on those systems, it is installed, fully operational, and in day-to-day use on those systems by a diverse set of DOE science projects. The directors of those centers have expressed their thanks to VACET for our assistance and efforts to help them achieve their primary mission, which is to enable science (see support letters in Section 11).

An important aspect of "production quality" software is the issue of software lifecycle management. The issues here are broad, and include "cradle-to-grave" concerns: initial regression testing, porting to different platforms, release engineering, ongoing support and maintenance. While addressing all of these issues is beyond the scope of this project due primarily to budgetary constraints and funding horizon, we have sensibly addressed many of these issues. Our team migrated the VisIt

[1]VisTrails (see `www.vistrails.org`) is a new scientific workflow and provenance management system. A peer-reviewed journal article about VisTrails won Best Paper Award at IEEE Visualization 2007.

code repository from “behind the fence” at LLNL to a publicly accessible location provided by the SciDAC Outreach Center, thereby opening the scope of community-wide software development from a limited few to a large, worldwide developer base.

Our initial software engineering and deployment strategy, as discussed in the project proposal, was to use VisIt as a deployment vehicle for new research. This approach has proven to be wildly successful for a number of reasons. First, we are able to effectively leverage several person-decades of NNSA investment in VisIt’s software and release engineering infrastructure: we, like our science stakeholders, were able to benefit from the “buy rather than build” approach to delivering results (see Sections 6 and 8). This infrastructure has enabled us to quickly deliver new results to science stakeholders, has provided a solid infrastructure base for conducting new research in petascale visualization (e.g., Section 3.1), and enabled us to deliver product that runs on all modern HPC platforms.

Another confirmation of the success of our original petascale visualization strategy is that VisIt is the first and only visualization application that is included as part of DOE’s *Joule Metric* (Section 5). That effort aims to provide quantitative evidence to OMB of increases in computational efficiency on petascale platforms, and is *prima facie* evidence that VACET is in the petascale visualization business. To this end, VACET is taking a leading role in helping the SciDAC program to achieve its mission of enabling science on petascale platforms.

#### 1.3.3 Field-Leading, Award-Winning Research

As one of the primary SciDAC objectives is to enable science, a vehicle for measuring success is the amount and quality of original scientific research performed by SciDAC projects. Our team has generated a vast amount of high-quality, award-winning research. This work has resulted in well over 100 scientific publications in the past 2.5 years (Section 9.1), including four Best Paper Award winners (three of which were at IEEE Visualization) and several Best Paper Award nominees at the annual IEEE Visualization conference.

Our broad research portfolio (Section 4) is driven by the needs of our science stakeholders and addresses many of the difficult problems in the field of visualization. Early in the project, we collected stakeholder needs and looked specifically for needs common across multiple areas, then focused R&D efforts into these areas to achieve the best possible impact for our R&D work. These research areas include large-data visualization; scalable visualization; flow visualization; feature detection, analysis, and tracking; multivariate and temporal visualization and analysis; statistical and uncertainty visualization. The vast number of peer-reviewed publications we have generated in all these areas is indicative of the quality of our team’s research efforts.

#### 1.3.4 Progress Towards Petascale

Our team has made substantive steps towards enabling petascale science as well as helping DOE and the SciDAC program achieve its objectives of effective use of petascale class machines. These efforts, discussed in more detail in Section 5, include: (1) providing a production quality visualization application, VisIt, as part of the *Joule Metric*, which provides quantitative measures of improvement in computational efficiency and scalability on DOE petascale platforms; (2) through leading-edge R&D activities, provide production quality software infrastructure that enables scientific knowledge discovery of data being produced by petascale applications on petascale machines; (3) working closely with science application teams as well as staff at DOE’s supercomputing facilities to solve problems that arise as we move towards the petascale, problems such as the design and implementation of effective data models, formats, and tuning/optimizing I/O libraries for use by

simulation codes on petascale machines as well as our visual data analysis software infrastructure.

### 1.3.5 Effective Use of ASCR Computing Resources

The VACET effort would not be possible without a close relationship with ASCR's major computing facilities and staff. VACET members include staff from two of the three DOE open computing facilities (NERSC/LBNL and LCF/ORNL). Having staff shared across both efforts enables us to rapidly deploy production quality, parallel capable visual data analysis software infrastructure at those facilities for use by a broad scientific community. Having shared staff is not a prerequisite – VACET software is also installed at the Argonne leadership facility – but it certainly helps. The value of VACET software and the relationship of the VACET relationship at those facilities is clearly spelled out by the directors of those centers in letters of support (Section 11).

> We greatly appreciate the support we have received from VACET in getting VisIt up and running at Argonne. We have had many INCITE users request VisIt for their visualization and analysis needs and VACET has been crucial to meeting thse needs. . . . All of these efforts have been important to the sucess of our mission. – *William Allcock, Argonne*

VACET has an allocation for storage and cycles at both NERSC/LBNL and LCF/ORNL (Section 10). We conduct R&D on those facilities, as well as apply our tools and techniques to "hero-sized" problems that would not be possible using desktop-class machines or departmental, mid-range clusters. Data intensive activities like visual data exploration and analysis demand a somewhat different architectural balance that compute codes: many of our algorithms are I/O-limited, and these facilities provide exceptional I/O capabilities for their scientific users. We look forward to expansion of those facilities in the future to include even more I/O bandwidth to high-speed storage systems as well as large, distributed memory GPU clusters.

### 1.3.6 Service and Outreach

VACET has participated in a vast number of workshops and delivered many tutorials, both at established events, such as the annual SciDAC program meeting, as well as on-demand for science communities, such as VisIt tutorial we gave to the fusion research community in Fall 2008. Our presence at workshops, like the Extreme Computing and the $E^3$ series last year, helps to guide the future of these important program development areas with respect to data understanding and exploration topics. A complete list of Tutorials and Workshops appears later in this document (Sections 9.3 and 9.6). We have also participated with the SciDAC Institute for Ultrascale Visualization in outreach activities. These activities include providing expert speakers at IUSV events.

VACET researchers have provided extensive service to the science community as reviewers, program committee, organizing committee, advisory board, conference chair, editorial boards, and so forth (Section 9.5) as well as dozens of invited talks (Section 9.2).

# 2 Summary of VACET Collaborations

## 2.1 Collaborations with Science Stakeholders

| Science Areat | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Accelerator | SciDAC SA: Community Petascale Project for Accelerator Science and Simulation (ComPASS), INCITE Awardee | Cameron Geddes, LBNL | • Develop and deploy new production quality, parallel capability to interactively explore particle-based datasets and to interactive track particles across time. Replaces a process that formerly required hours/days with one that requires only seconds/minutes.<br>• Develop and deploy novel visual interface to powerful data mining capability, release in production software.<br>• Develop novel unsupervised machine learning techniques to automatically locate and analyze beam particles in simulation data.<br>• Numerous joint publications in journals, conference proceedings.<br>• See Section 3.1. |
| Astrophysics | SciDAC SA: Computational Astrophyisics Consoritium: Supernovae, Gamma Ray Bursts, and Nucleosynthesis | Stan Woosley, UC Santa Cruz and John Bell, LBNL | • Provide project-wide visual data analysis software infrastructure.<br>• Provide assistance to team members to transition from older packages to VACET software.<br>• See Section 3.2. |

| Science Areat | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Astrophysics | INCITE: Multidimensional Simulations of Core Collapse Supernovae, PetaApps: Supernova Simulations with CHIMERA | Tony Mezzacappa, John Blondin, Steve Bruenn | • Develop and deploy petascale production turnkey visualization tool (VisIt) to interactively explore large time-varying datasets.<br>• Develop multidimensional visualization techniques for understanding complex radiation fields.<br>• Develop and deploy parallel streamlines algorithms for exploring magnetic fields.<br>• See Section 3.3 for more information about the work and Section 11 for Dr. Mezzacappa's testimonial. |
| Climate | SciDAC SA: Global Cloud Resolving Model, INCITE Awardee | Dave Randall, CSU | • Co-design and co-implementation of a high performance data model and I/O library for icosahedral meshes of highest-ever resolution cloud model.<br>• Develop parallel capable data loader for VisIt to load and process this new climate simulation data.<br>• Help PNNL and CSU teams transition to VisIt as primary visual data analysis software infrastructure.<br>• Provide assistance to GCRM in visualizing the massive amount of high-resolution data generated by the GCRM code.<br>• Tuning and optimization fo parallel I/O on `franklin.nersc.gov` to meet performance targets.<br>• See Section 3.5. |

| Science Areat | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Climate | SciDAC CET: Earth Systems Grid | Dean Williams, LLNL | • Design, develop, and deploy production-quality 3D/temporal-capable visual data analysis capability to the ESG and its many science stakeholders (e.g., CCSM Consortium).<br>• Design, develop, and deploy advanced comparative visual data analysis capabilities.<br>• See Section 3.4. |
| Climate | NOAA Geophysical Fluid Dynamics Library | Name of person at NOAA/GFDL | • Demonstrate visual data analysis of ultra-high resolution climate data runs – 0.25° atmospheric and oceanic grids – using VACET software. This result not possible with "standard" climate visualization applications.<br>• Develop custom data loaders to input `c360` and `cm2.4` simulation data into VisIt.<br>• Help GFDL team transition to VisIt for their large-scale visual data analysis needs.<br>• See Section 3.6. |
| Combustion | Interaction of Turbulence and Chemistry in Lean Premixed Laboratory Flames, INCITE Awardee | John Bell, LBNL | • Developed new capability for quantitative analysis of the relationship between turbulence and combustion efficiency. This work provides the first-ever view of this elusive relationship.<br>• Joint publications with stakeholders in high-impact *Combustion and Flame* journal.<br>• See Section 3.7. |

| Science Areat | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Combustion | High-Fidelity Simulations for Clean and Efficient Combustion of Alternative Fuels, INCITE Awardee | Jacqueline Chen, SNL | • Developed new capability for the quantitative analysis of combustion simulation data to better understand the genesis of extinction and reignition.<br>• New capability to help understand the role of transient autoignition.<br>• See Section 3.8. |
| Fusion | Framework Application for Core-Edge Transport Simulations (FACETS) | John Cary, Tech-X | • FACETS adopts VACET production-quality, parallel-capable visual data analysis software infrastructure for project-wide use in Fusion (and Accelerator) efforts.<br>• Researchers at Tech-X use VACET software to win a People's Choice Award at SciDAC 2008.<br>• Researchers at Tech-X using VACET software infrastructure as the base for customized, tailored visual data analysis applications.<br>• Using VACET visual data analysis software for visual data analysis of output of many different fusion simulation codes (NIMROD, FACETS, CSWIP).<br>• See Section 3.9. |

| Science Areat | Project Name | Collaborator | Summary of Accomplishments and Impact |
| --- | --- | --- | --- |
| Fusion | SciDAC Center for Gyrokinetic Particle Simulation of Turbulent Transport in Burning Plasmas (P. Diamond) and SciDAC Center for Plasma Edge Simulation (C.S. Chang) | Stephen Ethier, PPPL; Seung-Hoe Ku (NYU); Julian Cummings (Caltech) | • Develop new capability for rapid extraction and display of particles that meet investigation criteria, e.g., those that are undergoing "trapping and untrapping." Deploy in production-quality visual data analysis application.<br>• Design and implement data model strategy for indexing and organizing particle data in a form to support efficient visual data analysis parallel I/O and processing.<br>• Generate a functional prototype for multivariate subset selection and display.<br>• See Section 3.10. |
| Fusion | Center for Extended Magnetohydrodynamic Modeling (CEMM) (S. Jardin); Center for Simulation of Plasma Microturbulence (CSPM) (W. Nevins); Simulation of Wave Interactions with Magnetohydrodynamics (SWIM) | S. Kruger (Tech-X); J. Breslau (PPPL); W. Nevins (LLNL); D. Bachelor (ORNL) | • VACET delivered tutorials and workshops to help the fusion community, particularly at PPPL, migrate to VisIt as the community-wide visual data analysis software infrastructure.<br>• Develop Poincaré plot library, begin deployment in VisIt.<br>• Extensive outreach: tutorials (e.g., SciDAC 2007, 2008 meetings) and workshops (e.g., at PPPL for fusion researchers in Fall 2008) to help this community migrate to modern, production-quality visual data analysis software infrastructure.<br>• See Section 3.11. |

| Science Areat | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Turbulence | Simulations of Turbulent Flows with Strong Shocks and Density Variations | Eric Johnsen, Johann Larson (Stanford) | • Users of VisIt, including VACET-developed streamline algorithm<br>• Ported Silo library at Argonne to enable their runs there.<br>• See Section 11 for Dr. Johnsen's testimonial. |
| Turbulence | Reactor Thermal Hydraulics (INCITE runs of Nek5000) | Paul Fischer (Argonne) | • Developed pathline algorithm in VisIt for flow analysis<br>• Beneficiary of VisIt volume rendering improvements for several movies<br>• Modified VisIt's networking to enable remote collaborators to run client-server<br>• See Section 11 for Dr. Fischer's testimonial. |

Table 1: Summary of VACET stakeholder and collaborator interactions.

## 2.2 Technology and Outreach Collaborations

| Focus Area | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Mathematics | Applied Partial Differential Equations Center for Enabling Technology | Phil Colella, LBNL | • Extensive science needs assessment for AMR data knowledge discovery.<br>• Extensive software engineering to implement features needed to meet APDEC science needs.<br>• APDEC adopts VACET technology as its project-wide, production quality AMR visual data anlysis software infrastructure.<br>• APDEC saves money formerly spent on visualization software development and maintenance.<br>• Through partnership with APDEC, VACET visualization software adopted for use by researchers in fusion and accelerator modeling that use AMR-based codes, particularly Chombo.<br>• VACET are recognized as leaders in the field of AMR visualization by the computational community who perform AMR-based calculations as evidenced by adoption of VACET software for project-wide use.<br>• See Section 3.12. |
| Data Management | SciDAC Scientific Data Management Center for Enabling Technology | Arie Shoshani and Kesheng Wu, LBNL | • VACET and SDM Center researchers collaborate on numerous query-driven visualization R&D projects.<br>• VACET uses the FastBit software from the SDM center to accelerate query-driven visual data analysis projects. In one specific case, the duty cycle for scientific inquiry is reduced from hours to minutes/seconds. |

| Focus Area | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Outreach and Software Engineering Infrastructure | SciDAC Outreach Institute | David Skinner, LBNL/NERSC | • VACET makes heavy use of the software engineering infrastructure provided by the Outreach Center.<br>• VACET has given three tutorials at SciDAC conferences.<br>• See Section 11 for Dr. Skinner's testimonial. |
| Code Interoperabilty | SciDAC Interoperable Technologies for Advanced Petascale Simulations (ITAPS) | Lori Diachin, LLNL | • Mutual leverage on development in the VisIt project<br>• VACET's optimization of "subsetting code" directly benefits ITAPS.<br>• See Section 11 for Dr. Diachin's testimonial. |
| Visualization | Institute for Ultrascale Visualization | John Owens and Kwan-Liu Ma, UCD; Jian Huang, UTK | • VACET speakers at several IUSV workshops.<br>• Joint work on exploring use of GPUs to accelerate index/query operations: J. Owens (IUSV, UC Davis) and L. Gosink (VACET, UC Davis).<br>• Scalable data servers for large multivariate volume visualization: J. Huang (IUSV, UTK) and S. Ahern (VACET, ORNL). |

Table 2: Summary of VACET Technology and Outreach Collaborations.

### 2.3 Other Collaborations

| Focus Area | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Supercomputing Center | Argonne Advanced Leadership Computing Facility | Bill Allcock (Argonne) | • Ported, installed, and maintained VisIt on Argonne platforms.<br>• Modified VisIt network code to enable remote collaborators to run client-server.<br>• Ported Silo I/O library to Argonne machines.<br>• See Section 11 for Dr. Allcock's testimonial. |
| Supercomputing Center | LBNL National Energy Research Scientific Computing Center | Katherine Yelick and Francesca Verdier (LBNL) | • VACET applied for and received ERCAP allocation for cycles and storage during AY07, 08, and 09.<br>• VACET coordinates and collaborates activities with NERSC Analytics team to deliver production software and to provide assistance to numerous science stakeholders.<br>• VACET conducts many of its demanding research efforts on NERSC resources, including: scalability studies, visualization and analysis processing of massive datasets. Such resources are required because our projects demand resources far exceeding what is available on the desktop or in mid-range clusters.<br>• Ported, installed, and maintained VisIt on NERSC platforms.<br>• See letter of support in Section 11. |
| Supercomputing Center | ORNL National Center for Computational Sciences | Doug Kothe (ORNL) | • Ported, installed, and maintained VisIt on NCCS platforms.<br>• See letter of support in Section 11. |

| Focus Area | Project Name | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Visualization | VisIt project | Eric Brugger (LLNL) | • Took over significant responsibilities in the infrastructure of the VisIt project.<br>• Made software repository truly open (previously only available at LLNL).<br>• Established mailing lists for VisIt user community.<br>• Represents a significant collaboration between ASCR and the NNSA.<br>• See Section 11 for Mr. Brugger's testimonial. |

Table 3: Summary of VACET interactions with other collaborators.

> VACET's major development efforts have been great successes and have contributed greatly to the VisIt project as a whole. The parallel streamline algorithms they developed are cutting edge and filled what was previously a weak spot in VisIt. These algorithms have gone on to be used by customers outside the Office of Science . . . I believe that VACET development provides significant leverage to the NNSA efforts. – *Eric Brugger, LLNL*

# 3 Science Impact

In this section we give details of specific VACET activities that have had dramatic impact upon the science mission of our stakeholders. These range from new techniques for knowledge discovery to significant customer cost savings and increased efficiency.

> We depend on VisIt to visualize our results and appreciate the efforts of the VACET team, which enable us to use VisIt and are also keeping VisIt prepared for the petascale future. – *Eric Johnson, Stanford*

## 3.1 Accelerator: Advanced Accelerator Design – Laser-Wakefield Accelerators

### Background and Motivation

Laser wakefield simulations model the behavior of individual particles as well as the behavior of the plasma electric and magnetic fields. Output from these simulations can become quite large: today's

datasets, such as the ones we study here, can grow to be on the order of 200GB per timestep, with the simulation producing approximately 100 timesteps. The scientific challenge we help address in this study is first to quickly find particles that have undergone wakefield acceleration, then trace them through time to understand acceleration dynamics and perform both visual and quantitative analysis on the set of accelerated particles.

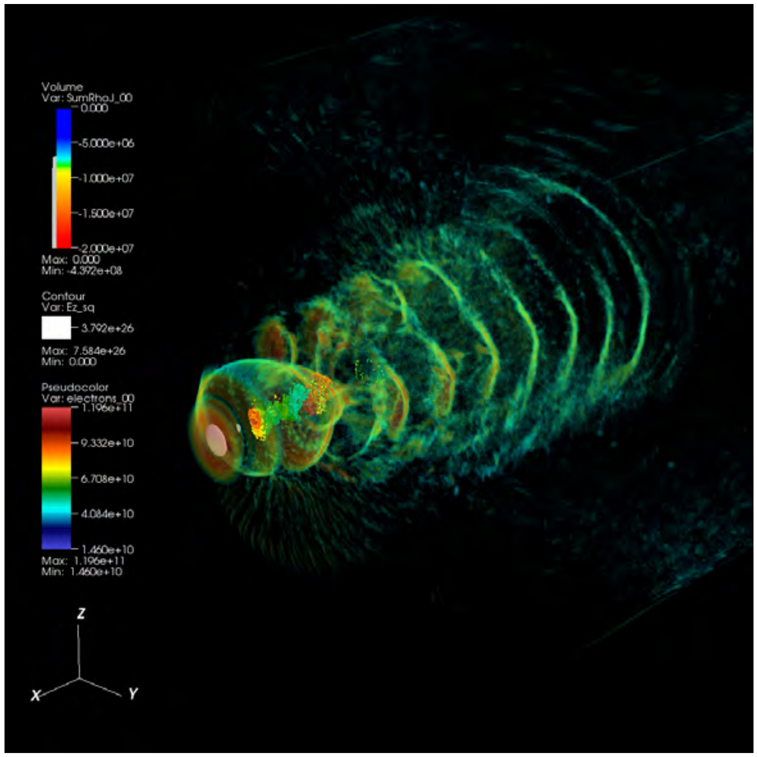



Figure 2: An intense laser pulse produces a density wake in a hydrogen plasma, shown by volume rendering in VisIt parallel visualizations of VORPAL simulations modeling LOASIS(LBNL) laser wakefield accelerator (LWFA) experiments. High energy particles, colored by their momentum, are shown here along with the wake. This combination helps facilitate understanding of how these experiments produced narrow energy spread bunches for the first time in an LWFA (red). The present simulations run on 3500 processors for 36 hours, and visualization of the 50 GB/time snapshot datasets runs on 32 `franklin` processors and requires tens of seconds per snapshot. Future simulation runs will increase these demands by orders of magnitude.

The primary objective of this project is to explore a novel technique for rapid data exploration and subsetting operations to support the types of investigatory work patterns in use by accelerator scientists. One scientific impact of our work is that we have vastly reduced the duty cycle in visual data exploration and mining. In the past, accelerator scientists would perform the "trace backwards" step using scripts that performed a search at each timestep for a set of particles. Runtimes for this operation were on the order of hours. Using our implementation, those runtimes are reduced from hours to seconds.

**Science Stakeholders and Collaborators**

The science stakeholders on this project, listed below, are part of the SciDAC Community Petascale Project for Accelerator Science and Simulation (ComPASS) project, and are INCITE awardees for cycles/storage at NERSC:

- Cameron G. R. Geddes and Estelle Cormier-Michel, LBNL.
- Peter Messmer, Tech-X Corporation.

In addition, we are collaborating with Kesheng Wu of the SciDAC Scientific Data Management Center on this project. Wu and the SDM Center have been instrumental in extending FastBit[2].

**Accomplishments**

1. **Data Modeling and Formats.** In order to apply FastBit-accelerated index/query operations to the laser wakefield simulation data, we used our HDF5-FastQuery technology (see `https://codeforge.lbl.gov/projects/h5part`), which provides a veneer API atop HDF5 and FastBit to implement simplified access to I/O and index/query capabilities. We also updated the VisIt plugin that loads such data.

2. **Effective Visual Information Display.** We developed a new technique for displaying parallel coordinates plots of large datasets using 2D histograms (having uniform- or adaptively-spaced bins.)

3. **Effective Query Interface.** We performed the software engineering in VisIt necessary to "link" the parallel coordinates plot, which provides sliders on each parallel coordinates axis to indicate subset selection, with the database plugin to extract the desired data subset from the larger data file. VisIt then passes the extracted subset from the file loader to downstream processing and visualization components.

4. **Visual Display of Query Results.** We developed a methodology for effectively presenting query results within the context of the larger dataset (see Figure 3). The idea is to provide two types of information: (1) to visually indicate the query results within the context of the original scientific datasets; and (2) to visually indicate statistics about the query result, which we accomplish with multiple, overlayed parallel coordinates plots.

5. **Accelerated Particle Tracking.** Once the conditions associated with particle acceleration are identified, we devised the means for quickly finding all such particles across the many timesteps of data produced by simulation. Results of primary/secondary acceleration, 3D visualization, and particle tracking are shown in Figure 4.

6. **Performance Experiment.** There are two main operations that comprise the HPC visual data analysis operation: (1) computing conditional 2D histograms for visual information display as well as the query interface; (2) particle tracking across the many timesteps produced by simulation. Our performance experiment uses a source dataset that is approximately 1.5TB in size, including the index data. We measure parallel performance and scalability of the histogram computation and particle tracking operations (see Figure 5).

[2]FastBit is software that implements high performance, compressed bitmap indexing. It won an R&D 100 award this past year. More information is located at `http://sdm.lbl.gov/fastbit`.

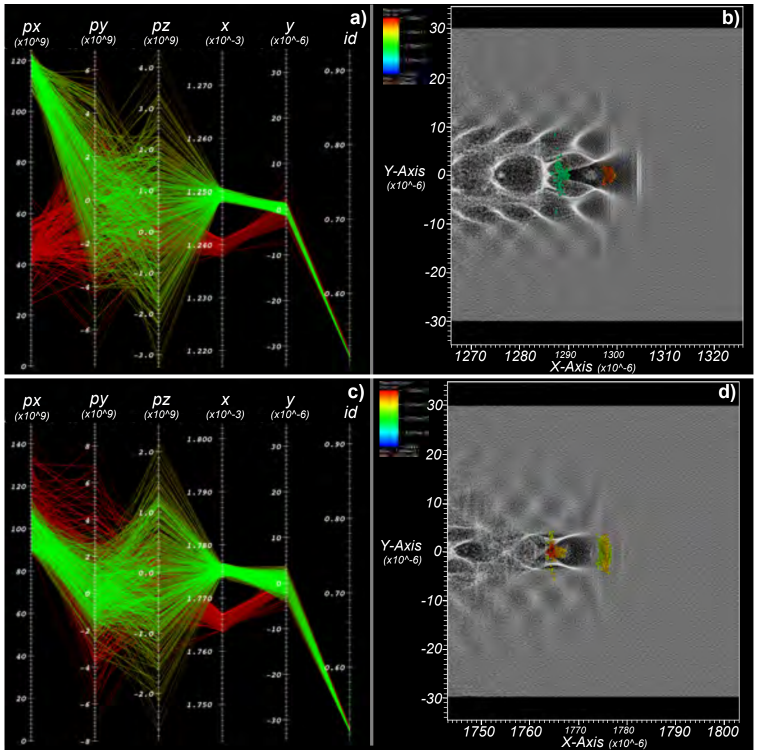



Figure 3: a) Parallel coordinates and b) pseudocolor plot of the beam at $t = 27$. Corresponding plots c,d) at $t = 37$. The context plot, shown in red, shows both beams selected by the user after applying a threshold of $px > 8.872 * 10^{10}$ at $t = 37$. The focus plot, shown in green, indicates the first beam that is following the laser pulse. In the pseudocolor plots b) and d), we show all particles in gray and the selected beams using spheres colored according to the particle's x-momentum, $px$. The focus beam is the rightmost bunch in these images. At timestep $t = 27$, the particles of the first beam (green in figure a) show much higher acceleration and a much lower energy spread (indicated via $px$) than the particles of the second beam. At later times, the lower momentum of the first beam indicates it has outrun the wave and moved into decelerating phase, e.g at timestep $t = 37$.

> Through a productive collaborative effort with experimental, theoretical and computational physicists from LBNL and Tech-X Corp., the VACET Center has had a very positive impact on our laser wakefield simulations. ...By allowing efficient analysis of large datasets, the VACET team has provided valuable support helping our team make progress towards petascale science, and is providing a valuable service to the scientific community. – *Cameron Geddes, LBNL*

### 3.2 Astrophysics: Computational Astrophysics Consortium

VACET's strategy for providing help to the computational astrophysics community focuses on teaming with the SciDAC Science Application entitled Computational Astrophysics Consortium (CAC): Supernovae, Gamma Ray Bursts and Nucleosynthesis. In a nutshell, that project combines computational and experimental science to increase an understanding of the evolution of the universe by focusing on supernovae, gamma ray bursts and nucleosynthesis.

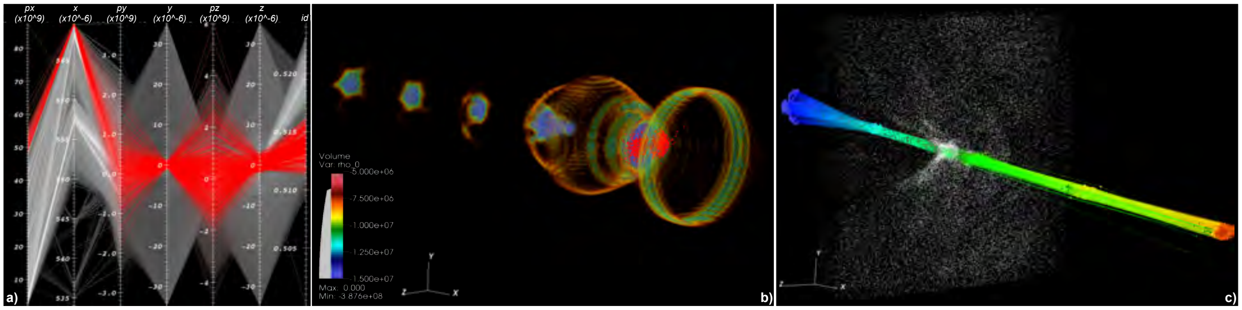



Figure 4: a) Parallel coordinates of timestep $t = 12$ of the 3D dataset. Context view (gray) shows particles selected with $px > 2 * 10^9$. The focus view (red) shows particles satisfying the condition ($px > 4.856 * 10^{10}$) && ($x > 5.649 * 10^{-4}$), which form a compact beam in the first wake period following the laser pulse. b) Volume rendering of the plasma density and the selected focus particles (red). c) Traces of the beam. We selected particles at timestep $t = 12$, then traced the particles back in time to timestep $t = 9$ when most of the selected particles entered the simulation window. We also traced the particles forward in time to timestep $t = 14$. Color indicates $px$. In addition to the traces and the position of the particles, we also show the context particles at timestep $t = 12$ in gray to illustrate where the original selection was performed. We can see that the selected particles are constantly accelerated over time (increase in $px$).

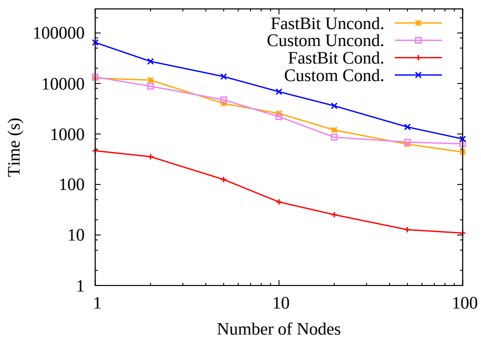
Our approach is to undertake efforts that provide visual and data analysis infrastructure that support both aspects of this large, ambitious astrophysics project. Our work with the simulation part of the community is outlined below in Section 3.2.1, and with the experimental community in Section 3.2.2. Both these communities are working together in the CAC: the experimental data provides validation of computational models, and the development of computational models helps guide how observations are performed.

### 3.2.1 Community Visualization Infrastructure

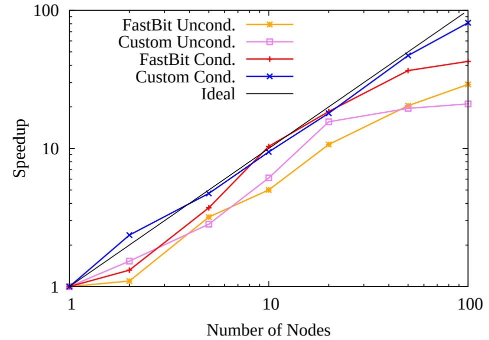
From a high level view, VACET's objective is to provide production-quality, parallel-capable visual data analysis software infrastructure for use by the SciDAC astrophysics community. To achieve that objective, we are working closely with "code owners" in the CAC to ensure that our primary production software, VisIt, meets their science-driven data understanding needs.

The CAC is relying on John Bell, LBNL, to perform code development of a low mach number, incompressible flow model of a Type Ia supernova that includes radiation transport. Our existing work with APDEC in providing production-quality AMR visualization software (see the VisIt and AMR Visualization two-pagers in Section 8) leaves us in very good position to quickly help this code team. At the April 2008 CAC All-Hands Meeting in Palo Alto, CA, John Bell told the CAC community that "VACET's VisIt is the tool of choice for performing visual data analysis of data from CASTRO" (CASTRO is his AMR-based code for supernova modeling). See Figure 6 for an example of this work.

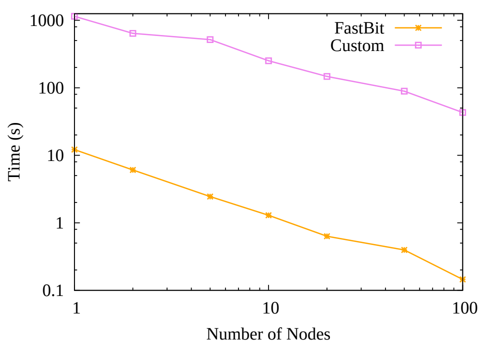
> Their team has been essential to our progress towards petascale science. VACET's codes and expertise have allowed us to visualize the complex datasets generated in our multidimensional simlations of stellar explosions. . . . Overall, VACET has had a positive impact on our science effort and is providing a valuable service to the scientific community. – *Stan Woosley, UC Santa Cruz*



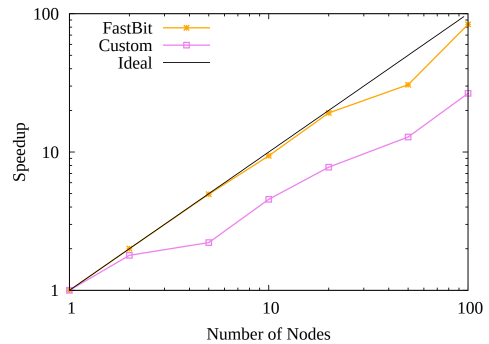

(a) Timings for parallel computation of histograms.




(b) Scalability of parallel histogram computation.




(c) Timings for parallel particle tracking.




(d) Scalability of parallel particle tracking.

Figure 5: These results show performance and scalability of parallel conditional histogram computation and temporal particle tracking. The science impact of this performance result is that the process of tracking particles undergoing acceleration across many timesteps of multi-terabyte datasets, a process that once took hours, now is accomplished in seconds. This capability is deployed in a production, quality visual data analysis application, and is readily applicable to other science domains.

VACET has also provided one-on-one support to other individuals in the CAC, namely Adam Burrows, in performing validation and exploration of data produced by the VULCAN code. Burrows is in the process of migrating from VULCAN to CASTRO for his work.

This project has a diverse and large set of needs that are documented in more detail on the VACET wiki at `http://www.sci.utah.edu/vacetwiki`. Near-term work plans focus on extending VisIt in a number of ways designed to meet a number of specific needs identified by the CCSE team at LBNL. As these needs are met, deployed in VisIt, and put into widespread use across CAC, we anticipate several things will happen. First, the needs list will grow longer and more focused on answering science-specific questions. These will be addressed in a way that achieves a good match between science priorities and VACET budget. Second, we expect our existing one-on-one relationships with CAC researcher to mature and become even more fruitful. Through these relationships, we anticipate significant science results and discoveries to emerge.

### 3.2.2 Spectra and Light Curve Visual Data Analysis

Complementary to the simulation thrust within CAC is an experimental thrust. The CAC has relationships with ongoing experiments: the Nearby Supernova Factory (SNF), the Joint Dark En-

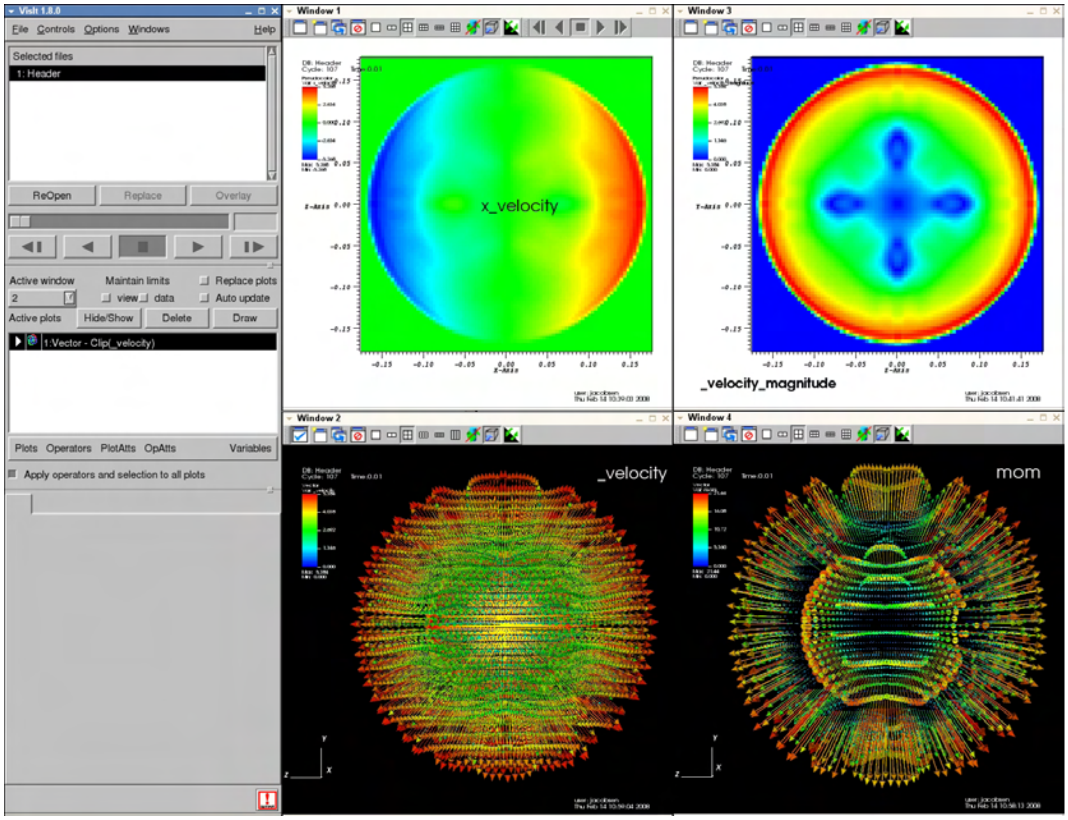



Figure 6: This image shows visual output of data generated by the CASTRO code, which is the AMR-based simulation in use by the CAC for modeling Type Ia Supernova explosions. The data was generated by Ann Almgren, LBNL, and the image created by Janet Jacobsen, LBNL.

ergy Mission (JDEM), the Supernova Acceleration Probe (SNAP), and the Large Synoptic Survey Telescope (LSST). The CAC uses an approach of validating simulation models against empirical observations, and these experiments all serve as sources of observations.

The visual data understanding needs of the CAC experimental group are documented on the VACET wiki, referenced above. In brief, those needs can be summarized as follows: (1) replace "chi-by-eye" dataset comparison techniques with those that are more rigorous, insightful, and scalable; (2) determine a good set of parameters for capturing the population of spectral and light curves obtained through observation for the purpose of classification, matching, comparison and so forth; (3) better tools for performing temporal analysis and evolution of spectra, light curves and their associated parameters.

For (1), VACET has several interrelated efforts in this space, e.g., comparative analysis and equivalence class functions. In addition, we have undertaken a separate project, called SpectraVis, that directly targets visual analysis and presentation of a large collection of spectra and light curves. An early prototype of that work, which was performed in collaboration with the Nearby Supernova Factory, provided a visual interface to browse a large collection of spectra and light curve data, and included some preliminary analysis. That work was well received by the science stakeholders. However, further development is currently in abeyance pending the arrival of a reference dataset from astrophysics collaborators in Europe.

While the SpectraVis work is on hold awaiting a validated collection of test data, we have focused on exploring new analysis and processing techniques that are useful for experimental astro-

physicists. We have developed a new technique for denoising noisy sky images, and experimented with applying fundamental analysis algorithms to particle accelerator data. We have also submitted research papers on the subject of effective data presentation and exploration methods useful to the experimental astrophysics community. In the long run, the intent is to simplify scientific knowledge discovery for the astrophysics community: this community is faced with trying to gain scientific insight from data from diverse sources: collections of experimental/observed data and simulation results.

### 3.3 Astrophysics: Simulation of Core-collapse Supernovae with CHIMERA and GenASIS

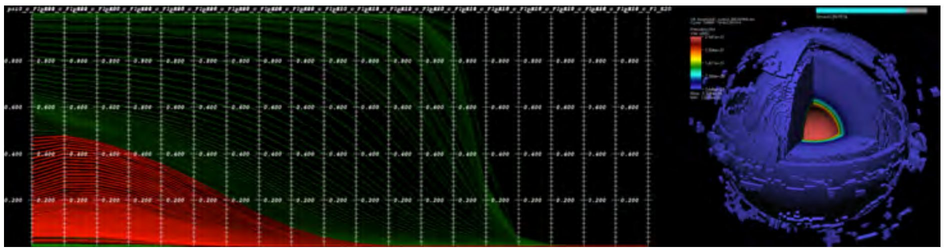


Figure 7: Multidimensional parallel coordinates exploration of the 20-energy group neutrino radiation field in a large supernova simulation. Increasing neutrino occupation highlighted in red. The multidimensional analysis on the left is directly linked to a traditional spatial visualization on the right. Visualization: Sean Ahern (Oak Ridge National Laboratory). Simulation: Steve Bruenn (FAU), Bronson Messer, Tony Mezzacappa, Raph Hix (Oak Ridge National Laboratory)

Core-collapse supernovae are arguably the most important links in our chain of origin from the Big Bang to the present day. They are the dominant source of elements in the universe, producing all of the elements in the periodic table between oxygen and iron and half of those heavier than iron. Ascertaining the explosion mechanism for core-collapse supernova is one of the most important unsolved problems in astrophysics. Dr. Tony Mezzacappa at Oak Ridge National Laboratory is the head of a multi-institutional effort to simulate the complex physics of core-collapse supernovae. The primary simulation codes used are CHIMERA and GenASIS, used in both the DOE INCITE and the NSF PetaApps programs.

In support of Dr. Mezzacappa's petascale computational effort, the VACET team has worked to assist them with their scientific objectives. As their data has grown in size over the past several years, one early challenge was to provide support for the VisIt visualization system on the Oak Ridge National Laboratory Jaguar system. This was a significant porting effort, as the programming and filesystem environment was not conducive to rapid development of interactive visualization systems. Nevertheless, staff were able to port VisIt and provide support for the astrophysics team in loading their data in parallel. This porting effort has been reproduced at LBNL/NERSC. The production visualization support, provided in part by VACET staff, has allowed the astrophysics team to move entirely to VisIt as their primary visualization tool.

The CHIMERA simulation code has a very complex data model. Not only does it simulate time-dependent magnetohydrodynamics, but it also incorporates anisotropic multifrequency neutrino radiation, taking into account all three "flavors" of neutrinos (electron, $\mu$, and $\tau$). This accounts for an 8-dimensional data space for the neutrino radiation alone, much more complex than many other simulation domains. Researchers also wish to understand the correlation between the radiation

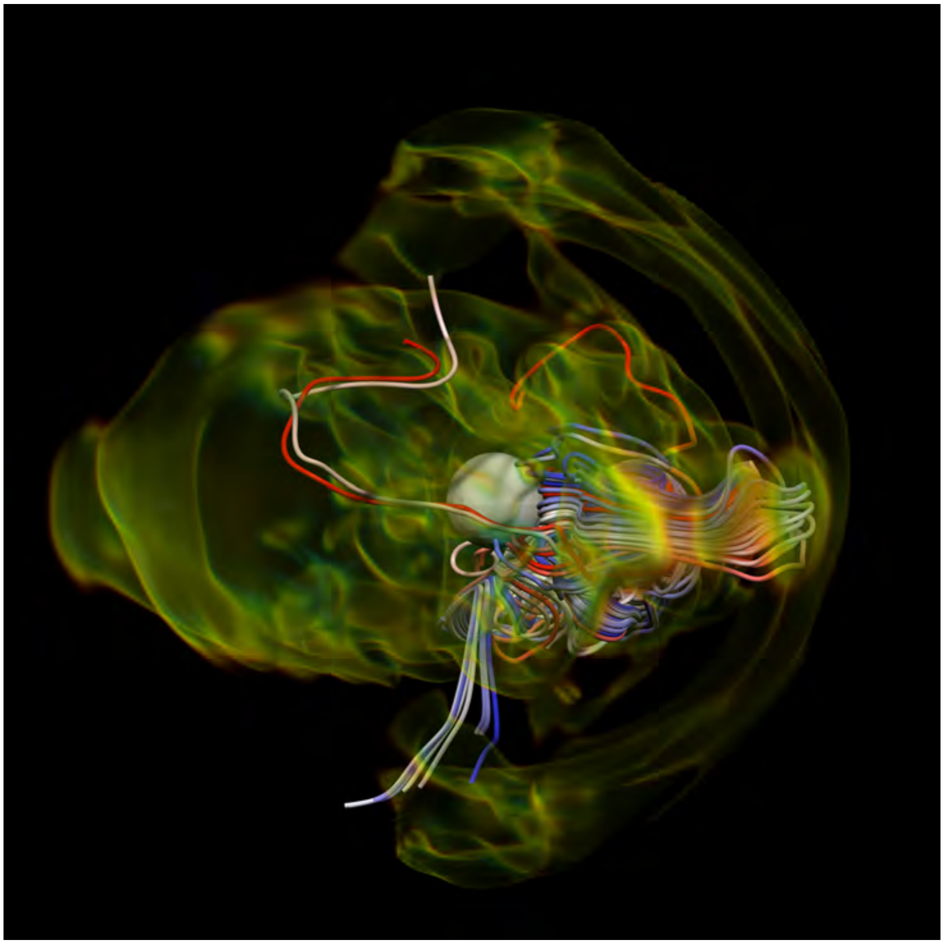


Figure 8: The collapse of massive star's core results in the formation of an outgoing spherical shock wave that eventually disrupts the entire star, giving rise to a supernova. Along the way the shock temporarily stalls and experiences the "stationary accretion shock instability" (SASI), which causes large deviations from spherical symmetry. This appears to be important to the supernova explosion mechanism, and may be responsible for spinning up the collapsed core – a nascent neutron star – into a pulsar. This image shows an exploratory view of a simulation run to ascertain the extent to which the SASI may generate magnetic fields: a volume rendering shows the fluid speed, and a sampling of fluid streamlines is colored by magnetic field strength. The simulation was run on Jaguar at NCCS with GenASiS, a multi-physics code under development for the simulation of astrophysical systems involving nuclear matter. Image credit: Visualization: Dave Pugmire (Oak Ridge National Laboratory). Simulation: Eirik Endeve, Christian Cardall, and Reuben Budiardja (Oak Ridge National Laboratory and University of Tennessee, Knoxville)

and the vector-valued magnetic field, as well as their relationships to scalar values such as entropy, pressure, and temperature.

To get a handle on the neutrino radiation field, VACET staff members adapted a multidimensional analysis technique from "infovis," that of parallel coordinates, and have deployed it within VisIt. Staff developed a data parallel binning method that allows parallel coordinates visualizations to scale to any size dataset. This method has proven particularly effective for understanding multifrequency radiation, as it has allowed Dr. Mezzacappa's team to directly visualize, for the first time, the neutrino radiation spectrum across the whole problem. See figure 7.

In addition to the radiation field, understanding the interactions between the magnetic field and

other attributes of the data model is important. VACET staff members have recently developed and deployed several parallel streamlines algorithms that allow streamlines to be traced in a scalable around the way across multiblock simulation meshes. Staff have employed this new method to explore features of the magnetic field as it evolves around the supernova shock front and the proto-neutron star. See figure 8.

ORNL visualization staff, in collaboration with VACET, have generated many visualizations for the astrophysics team and have assisted them in learning to use VisIt to perform their own visual data analysis. VisIt's ability to handle remote visualization using parallel computational resources allows the teams to run a client on their local workstation and do all analysis and visualization on institutional servers with fast access to their simulation data.

> Through a productive collaborative effort with our science team, the VACET Center has had a positive impact on our astrophysics project. . . . the VACET team has ported the VisIt visualization system to the Cray XT architecture, allowing us to analyze and visualize the largest datasets we can generate. In addition, they have developed multidimensionsl visualization techniques that have, for the frist time, allowed members of my team to directly view spatially varying multifrequency neutrino energy distributions, critical to the understanding of supernova shock acceleration. . . . the VACET team has provided valuable support helping my team make progress towards petascale science and is providing a valuable serivce to the scientific community. – *Anthony Mezzacappa, ORNL*

### 3.4 Climate: Earth Systems Grid

The most advanced climate modeling systems seek to enable a new deeper understanding of the dynamics of global carbon cycle, atmospheric chemistry, land and ocean ecological processes and their coupling with climate. These new, advanced climate models will enable pursuing reliable answers to fundamental questions related to climate variability and global change at time scales ranging from decades to centuries. In this effort, VACET will work in close collaboration with the Earth Science Grid and provide new advanced data analysis and visualization tools to the CCSM Consortium and the climate modeling community in general. One target will be the deployment of a first set of tools by the end of FY08, in time to facilitate the analysis of data for the 5th assessment report of the Intergovernmental Panel on Climate Change (IPCC)[3]. VACET is committed to supporting the needs of the Community Climate System Model (CCSM) Consortium[4] in collaboration with the Earth System Grid[5]. Our climate stakeholders include:

- Community Climate System Model Consortium, Primary contacts are John Drake (ORNL) and Phil Jones (LANL). Drake and Jones are co-PIs of the SciDAC Science Application entitled "Modeling the Earth System."
- Earth System Grid. Primary contact is Dean Williams, PI of the SciDAC Center for Enabling Technology entitled "Sharing a World of Data – Scaling the Earth Systems Grid to Petascale Data to enable faster, easier sharing of climate change research data."

[3]http://www.ipcc.ch/
[4]http://www.cgd.ucar.edu/csm/
[5]http://www.earthsystemgrid.org/

- Primary target software tools: CDAT, VCDAT. New VACET capabilities will be layered atop the well established CDAT/VCDAT software infrastructure.

As this effort has a long time horizon, a summary of our objectives in this space are as follows: (1) deploy advanced visualization capabilities into the CDAT tool and create a clear path for similar integration in other tools; (2) extend the visualization software to incorporate domain specific requirements, data formats, and vector field visualization; (3) support time-dependent and cross-dataset comparison, visualization and analysis; (4) develop new analytic capabilities for climate data (first deployed into CDAT/VCDAT); (5) integrate with the VisTrails framework for tracking and logging of internal state and provenance information; (6) develop a visualization and data analysis scenario for understanding of complex coupled phenomena such as the multi-scale dynamics the complete carbon cycle on earth.

To date, the bulk of our work has focused on (1), which aims to extend the fundamental analysis and 2D visualization capabilities of the Climate Data Analysis Toolkit (CDAT), provided by ESG to its stakeholder community, with a baseline set of 3D visual data analysis tools. The basic idea is that by adding capability to CDAT, we will have a faster and broader positive impact on the climate community than if we ask them to migrate to a separate application (e.g., VisIt) for doing 3D/4D visual data analysis work.

> Our collaboration with the researchers of the VACET has been very useful and we are committed to continue this partnership in the future. The tools provided by the VACET team will prove very useful for climate scientists and are allowing us to make progress towards an infrastructure for the exploration and understanding of petascale scientific data. – *Dean Williams, LLNL.*

In this work, we have focused on software and release engineering issues that target fine-grained milestones associated with providing production-quality software infrastructure that meets objective (1). In a nutshell, we are adapting and extending software infrastructure for use in the CDAT environment, as well as performing fundamental unit testing. Results of this work are shown in Figures 9 and 10.

We are exploring alternative scene layouts for data exploration and movie creation. For example, the image below (Figure 10) shows a possible layout including both embedded 3D data and slices on the side. This type of layout can increase both the realistic presentation of the information and the effective communication scientific the results. After completing our discussion with John Drake, Jamison Daniel, and the rest of the climate team at ORNL, we will improve the layout and content of the scene and generate a complete animation via scripting embeddable in CDAT. This capability will allow ESG to give the script to their users so they can produce similar presentations by simply changing the data source in the python script.

Next steps will take a more science-centric focus. The idea will be to team up with individual climate scientists and begin to apply the new 3D visual data analysis capabilities to specific science problems. The intent here is to better understand strengths and shortcomings of applying staple 3D visual data analysis techniques (presented via CDAT) to specific scientific inquiries. We expect several different types of results from those interactions. First, we anticipate those interactions will provide priorities for next steps: will we better understand what additional 3D visual data analysis capabilities are needed by this community, and we will have a better understanding of scientific priority, which in turn will prioritize our work. Second, our work will have a direct, positive impact on ongoing climate science concerns, which is one of the primary VACET missions.

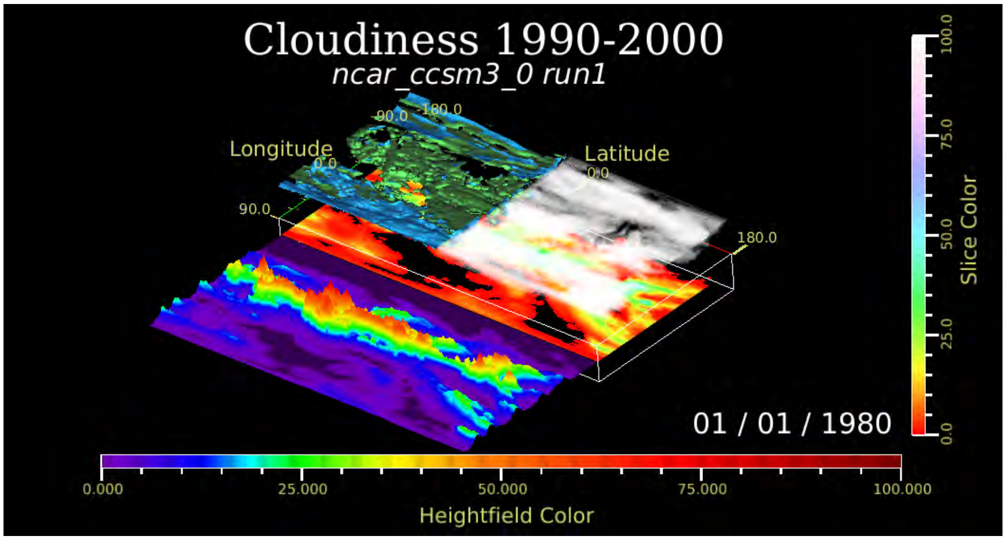



Figure 9: This image represents a new capability – a prototype integration of advanced visualization capabilities into the ESG-sponsored Climate Data Analysis Toolkit. This result proves the feasibility of the VACET approach to layering new capabilities onto high-quality software for use by a broad community.

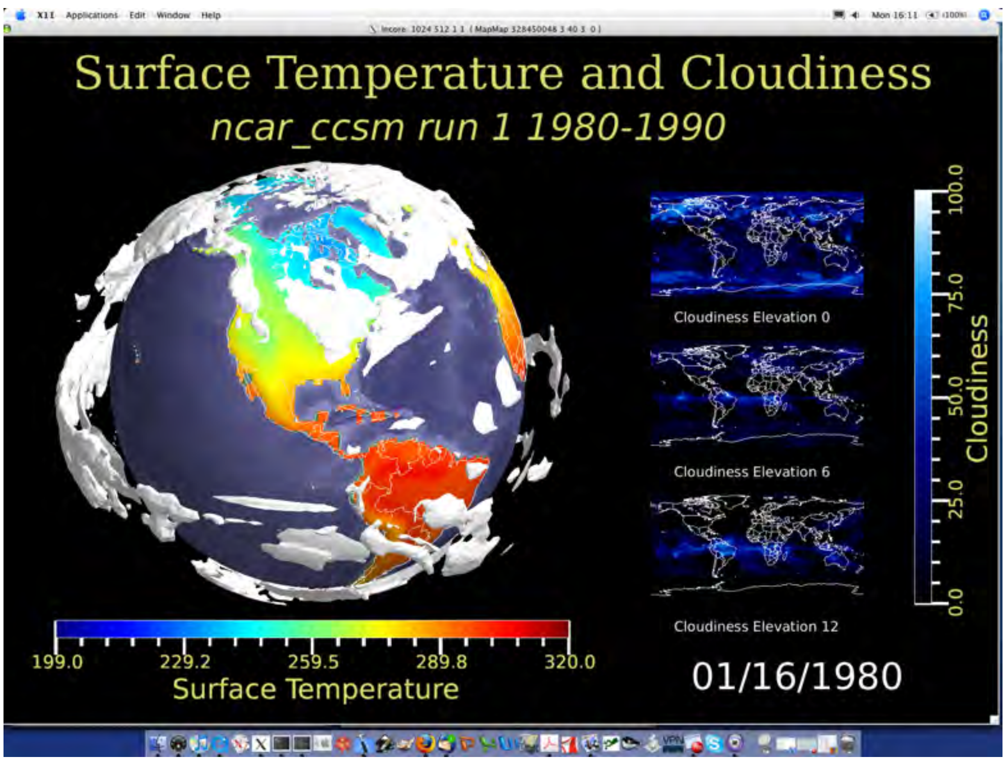



Figure 10: This image shows use of both traditional 2D and new 3D visualization capabilities within CDAT. This type of layout can increase both the realistic presentation of the information and the effective communication of the scientific results.

### 3.5 Climate: Global Cloud Resolving Model

In support of the SciDAC Science Application "Global Cloud Resolving Model" led by Dave Randall, CSU, the VACET team has collaborated with the NERSC Analytics team in support of several interrelated objectives. These include: contributing to tuning and optimization studies aimed at helping the Randall team meet minimum I/O parallel I/O performance levels; helping to design, implement, and test a parallel I/O data model and file format suitable for use by the code team and for subsequent analysis; helping GCRM team with challenging visualization problems that result from the unique geodesic grid and the massive amount of data produced by the GCRM code.

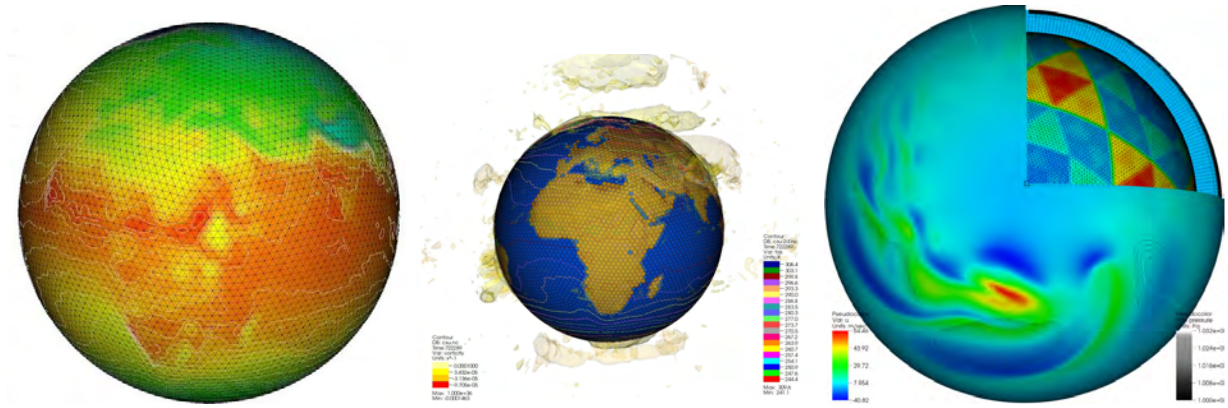
Dave Randall's SciDAC and INCITE projects propose to run GCRM simulations at the 3km resolution. At that resolution, a 24-hour run on 30,000 nodes will generate 10TB of data. The target ratio of time for compute is 95%, with 5% time allocated to doing IO. This target ratio amounts to a requirement of 2GB/s sustained collective I/O performance. Prior to this effort, sample benchmarking code obtained a collective IO performance of 0.3-0.8 GB/s on `franklin` (NERSC).

**Parallel I/O Tuning and Optimization.** NERSC Analytics staff in collaboration with VACET and other NERSC systems staff, led a detailed exploration into the source of the poor collective I/O performance on the Cray XT4. A number of issues were explored at various levels of the system/software stack. The ranged from making sure that the right collective NetCDF calls were being made, to MPI-IO optimizations, to block/buffer size allocations on individual nodes, Lustre striping specific hints, I/O hardware on the OSTs and so on. Some issues were identified with the closed-source MPI-IO implementation provided by Cray, and the vendor is working on resolving the problem. In the meanwhile, a recommendation on ideal settings for various system parameters has been communicated to our CSU/PNNL collaborators. Aggressive tuning by NERSC personnel indicates that collective I/O performance above 2GB/s is possible on `franklin`, but more work needs to be done with Cray to further optimize the IO performance.

> Through a productive collaborative effort with our teams, the VACET Center has had a positive impact on our science projects. . . . In combination with the expert help on IO profiling and troubleshooting on Franklin, VACET has had a positive impact on our science efforts, and is providing a valuable service to the scientific community. – *Karen Schuchardt, PNNL*

**Parallel file format for GCRM.** NERSC Analytics staff in collaboration with VACET worked with Karen Schuchardt, et al. (PNNL) to draft a data model and parallel file format for GCRM simulations. This effort targeted towards supporting all variants of 2D/3D meshes and data variables. The file format and data model are addressing the problem of doing massively parallel, scalable I/O on 30K+ cores to a single shared file. The existing file format is self-describing to ensure compatibility with production quality visualization and analytics tools such as VisIt. The current data model is being implemented using the CF conventions in NetCDF 3.0. Given the 32-bit file limitations in NetCDF, it is anticipated that final implementation would be done in NetCDF4.0 or HDF5.

**Visualization and Analytics.** GCRM simulations generate an unprecedented amount of data. It is estimated that the geodesic grid will have billions of points and tens of variables at each mesh point, resulting in TBs of data. Conventional tools are simply incapable of handling datasets of this size.



(a) A "first-light" image of a 2D icosahedral mesh colored by surface temperature.

(b) Another "first-light" image of a 3D icosahedral mesh, land cover, surface temperature, and atmospheric vorticity.

(c) VisIt plot of 3D icosahedral meshes and variables defined on cell centers, corners, and edges.

Figure 11: Visualization of atmospheric GCRM data using the VisIt visualization system.

We wrote modules to directly import the GCRM data into VisIt. This avoids costly data conversion operations and leverages VisIt's large set of visualization and analysis capabilities. The import modules support 2D and 3D meshes. Most of the GCRM variable types are supported by the VisIt module. Figure 11 demonstrates visualization of wind velocity and pressure on the 3D geodesic grid, with a cutaway highlighting the underlying meshes.

> Our in-house tools are simply incapable of handling the large data volume generated by our GCRM simulations. Thanks to the parallel capabilities and rich feature set of VisIt, we are now in a position to visualize and analyze our data and gain valuable insights. – *David Randall, CSU*

NERSC Analytics staff, in collaboration with VACET, have generated numerous visualizations for the GCRM team and have helped them to learn to use VisIt to perform their own visual data analysis. Due to VisIt's ability to support remote/distributed and parallel visual data analysis enables the GCRM team to run a client on their laptop, and stream imagery from a back-end servers running on at centrally located facilities with fast access to the simulation data.

### 3.6 Climate: NOAA-GFDL

In work conducted in collaboration with the NERSC Analytics team, VACET researchers enabled visual data analysis of ultra-high resolution climate data, and helped a climate research team transition to production-quality visual data analysis software infrastructure capable of handling their large data requirements and meeting their scientific knowledge discovery needs. The information below is drawn from a news release posted at NERSC[6].

Two teams of researchers pushing the limits of climate and weather simulations achieved noteworthy results running their codes on Franklin. The climate scientists successfully ran experimental

[6] `http://www.nersc.gov/news/annual_reports/annrep07/html/nersc_center/early_results.html`

simulations with resolutions much higher than in widely used codes. And another team of researchers set a speed performance record for a U.S. weather model, running on 12,090 of Franklin's processors.

> VACET researchers have deployed production-quality, parallel-capable software at DOE supercomputing centers and at "end stations" that meets the visual data analysis challenges resulting from increasing spatiotemporal climate model resolution that accompanies the transition to petascale-class platforms.

At the DOE's behest, scientists from the National Oceanic and Atmospheric Administration's Geophysical Fluid Dynamics Laboratory (GFDL) proposed a set of experiments using climate models with resolutions many times higher than those in the standard models, such as those used by the Intergovernmental Panel on Climate Change (IPCC). The high-resolution models offer not only a closer look at physical elements of the climate, such as tropical storms, but they also enable researchers to conduct a more in-depth analysis of climate change as higher-resolution phenomena in the ocean and atmosphere are resolved.

The research and development work needed to support this effort includes:

- Development of custom data loaders for VisIt. This work results in a time/cost savings by eliminating a costly data conversion step: prior to this work, it was necessary to perform a data conversion step from the format generated by the climate codes into a format suitable for loading into VisIt.
- Joint visualization work with the stakeholders. Here, the climate scientists were interested in focusing on the most interesting and significant phenomena: formation of tropical storms and ocean eddies.
- We installed VisIt on machines at GFDL and provided tutorials, both written (web-based) and in-person, to help GFDL researchers successfully adopt VisIt for the large-scale climate data analysis needs.
- One of the animations from this work won a *People's Choice Award* at the SciDAC 2008 program meeting.

### 3.7 Combustion: Interaction of Turbulence and Chemistry in Lean Premixed Laboratory Flames

Our team has been working with John Bell (LBNL) and researchers at LBNL's Center for Computational Sciences and Engineering (CCSE) to perform research, development, and application of topological analysis techniques to provide new insights into the combustion process. Bell's group works with large AMR 3D time-varying combustion simulation datasets, and needs analysis tools to aid their understanding. They are interested in formation and evolution of "extinction zones." They may be able to provide us with a boolean volume of "burning" and "non-burning" regions, and we will perform a detailed analysis of the structural characteristics of the network of low temperature regions. They are also interested in using topological methods to identify those zones with topology-based methods. To this end we will use topological tools both feature characterization and for robust time tracking.

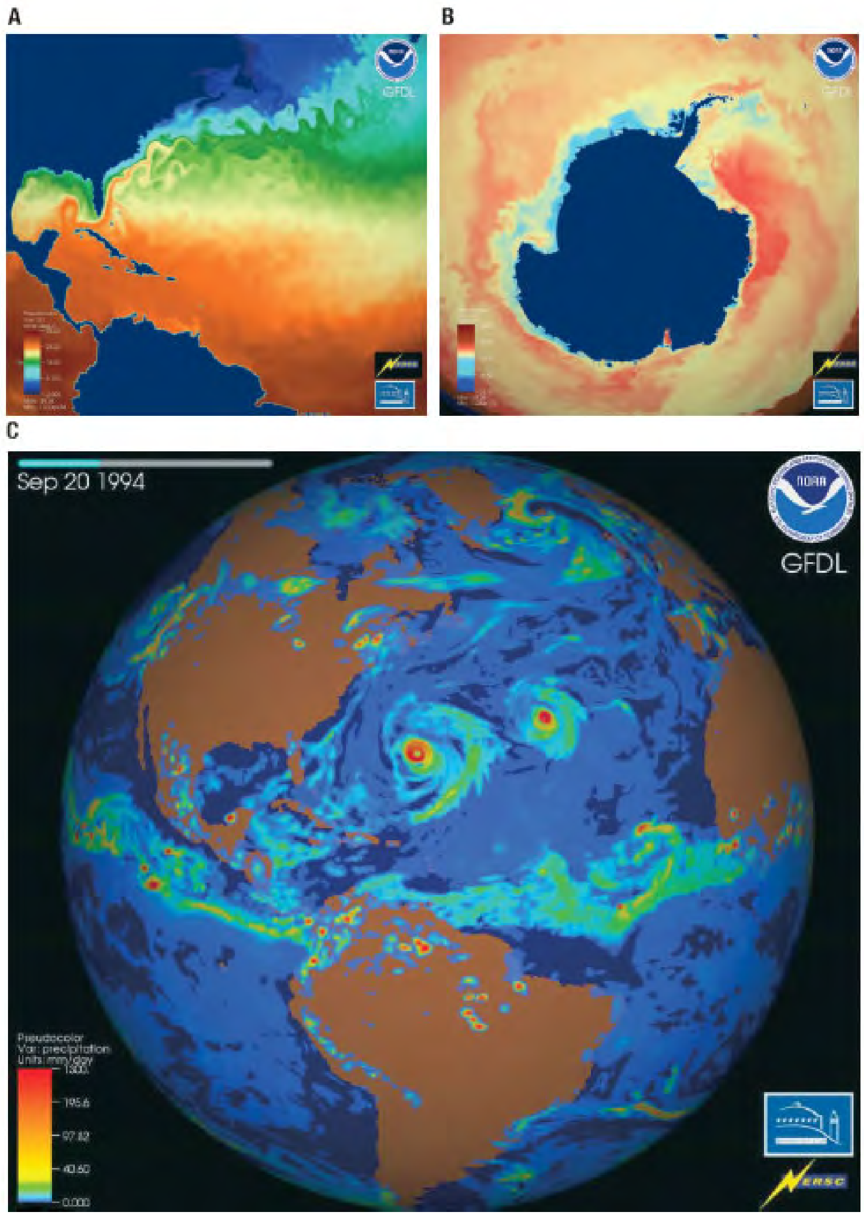



Figure 12: Visualizations of high-resolution climate datasets: (a) sea surface temperatures for the North Atlantic Gulf Stream; (b) salinity of the Southern Ocean; (c) hurricanes forming in the Atlantic Ocean. (Datasets provided by Chris Kerr, NOAA/GFDL; visualizations by Prabhat, NERSC/LBNL/VACET.)

We have been able for the first time to achieve a quantitative analysis of the impact of relative turbulence intensity on cellular burning structures in lean premixed hydrogen flames. – *John Bell, LBNL*

**Accomplishments**

We have expanded and finalized the topology based analysis of three simulations of lean hydrogen flames under different levels of turbulence. Understanding combustion processes over a broad range of operational regimes is of great interest for a variety of applications such as engine or power plant design. To this end, there has been considerable recent interest in the development of premixed burners capable of stably burning ultra-lean hydrogen-air fuel mixtures. Such burners could, for example, be used as one component of a clean-coal power plant utilizing hydrogen extracted from coal gasification. Lean premixed systems are subject to a variety of hydrodynamic and combustion instabilities that render practical flame stabilization, and traditional approaches to flame analysis, extremely difficult. The flames burn in a cellular mode that is highly nonuniform, time-dependent, and difficult to characterize. The stake holders are interested in understanding how many independent burning cells exist at any one time, their areas, and evolution over time. In particular the focus is on comparing the flames at different turbulence levels.

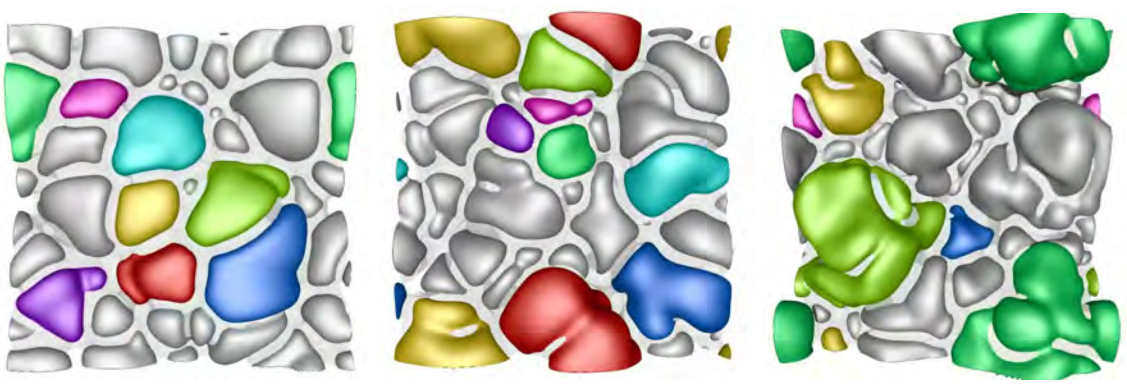


Figure 13: These images show the results of topological analysis that segment data into regions of combustion. The source data are no turbulence (left image), weak turbulence (middle), and strong turbulence (right).

We track all burning cells in all three simulations and create tracking graphs that encode the temporal evolution of all cells (Figure 15. The graph provides the stakeholders with in-depth information about how individual cells behave over time. The graph is used in multiple ways: First, by propagating colors along edges of the graph we can render color coded animations to show the evolution of certain cells. Second, the graph can be used to study interaction between cells on a individual scale and is also used for debug the segmentation algorithms. Finally, we use the graph to compare different spatial and temporal resolutions of each simulation to tease out similarities and differences.

Figure 15 shows the results of one such analysis, where we see a small portion of the tracking graph for the turbulence free case. Round nodes indicate burning cells segmented from an actual flame surfaces using the Morse complex. The numbers inside the nodes and along the branches indicate the identifier assigned to this particular cell in the segmentation. The diamond shapes represent topological events in between time steps and thus have no identifier attached. To provide

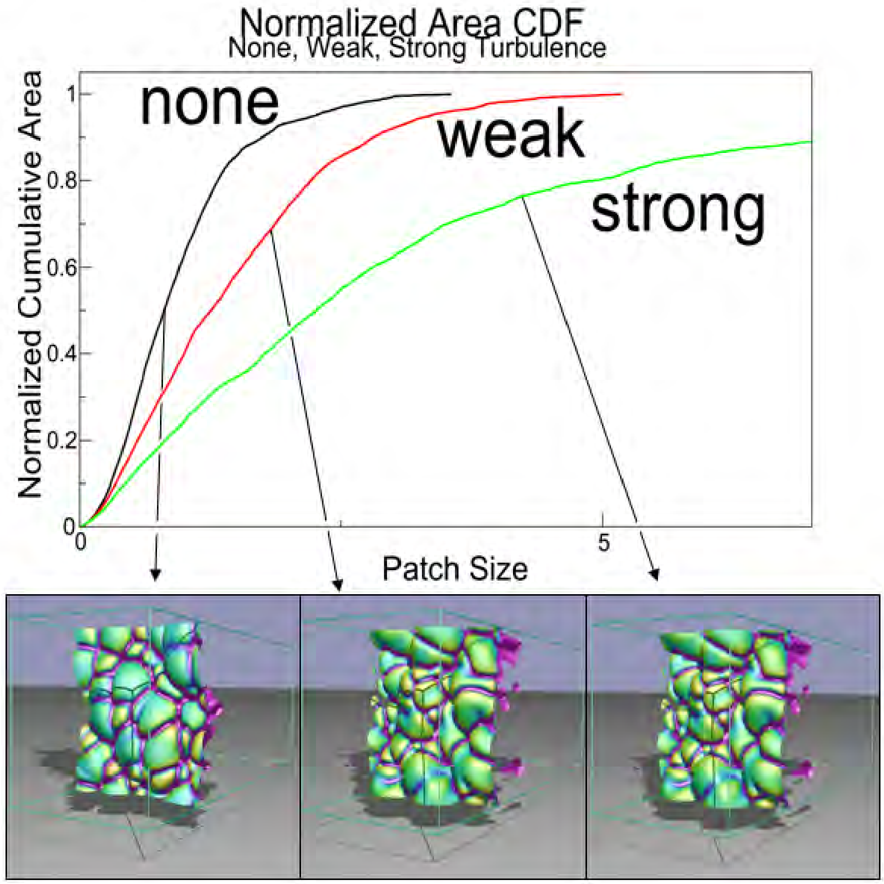



Figure 14: The topological analysis provides the means for quantitative comparison of combustion region attributes in the presence of no, weak, and strong turbulence.

better navigation we use the space inside the diamonds to indicate the simulation time. Red/green nodes and diamonds signal a merge/split event and turquoise structures a birth or death event. Using the renderings above and below the graph one can follow the event chain.

Between time 550 and 552.5 the purple cell 4 merges with cell 18 via a small connection across the lower (periodic) boundary. Subsequently, what used to be cell 18 develops a connection with cell 25 resulting at time 555 in the three purple areas all connected via small bridges to a single cell. At time 557.5 the cell formerly labeled 25 has split leaving a small portion attached to cell 4 and creating a new cell 26. Finally, at time 560 the remains of cells formerly labeled 18 and 25 have split off cell 4 forming a separate cell 19.

**Future Work**

**Near Term**

- Develop new streaming segmentation algorithm capable of dealing with three-dimensional features rather than features on surfaces.
- Implement the new segmentation algorithm.

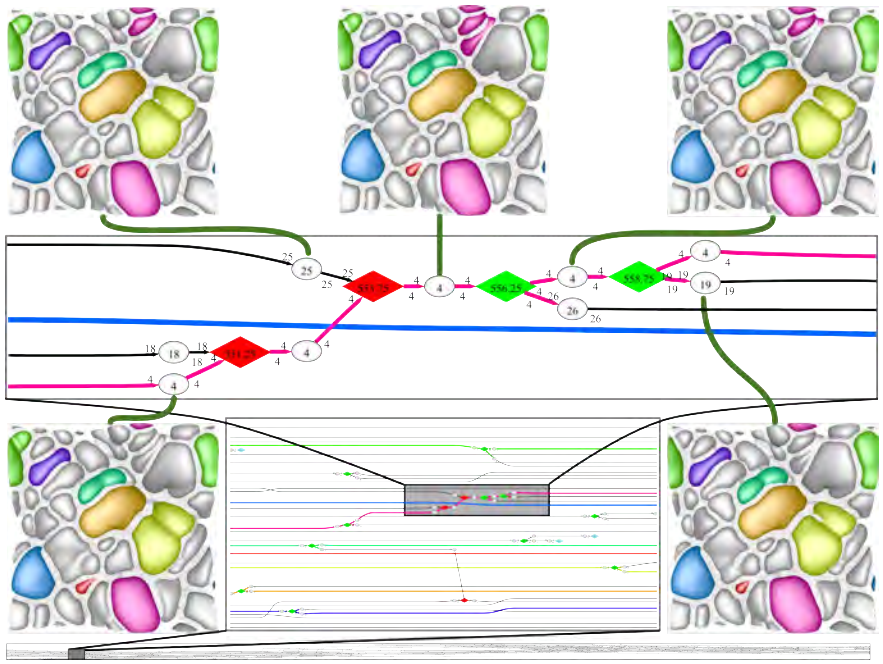


Figure 15: Detailed graph analysis applied to the turbulence-free dataset. It provides detailed information about the evolution of burning cells/regions, including split/merge operations that would not be possible with traditional analysis methods.

- Expand two-dimensional analysis to analyze the genus of non-burning regions.

**Longer Term**

- Adapt the tracking algorithm to three-dimensional features.
- Apply the new pipeline to a larger, more complicated data of interest to the stakeholders.
- Compare results from two-dimensional and three-dimensional techniques.

## 3.8 Combustion: High-Fidelity Simulations for Clean and Efficient Combustion of Alternative Fuels

Our stakeholder, Jacqueline Chen at SNL-CA, has the need of being able to perform quantitative analysis of data produced by DNS combustion simulations to aid in the understanding of the combustion process. Our approach is to perform topological characterization of combustion features in DNS-produced data. The work has a particular focus on developing new insight about the genesis and evolution of "extinction pockets." Understanding how extinction and reignition happens in premixed hydrogen combustion has the potential of allowing better design of engines and power plants. We focus on the use of robust topological methods for segmentation and tracking of regions of high scalar dissipation rate that provide a good first order approximation of true extinction regions.

**Accomplishments**

**Multi-scale Segmentation of Dissipation Rate.**

We completed the computation of multi-scale segmentation of regions high scalar dissipation rate. Worked with stakeholders at Sandia to develop a new S3D library and interface layer to run segmentation code. It features: (1) serial streaming mode with minimal memory and hardware requirements; (2) parallel mode coupled to S3D with all derived variables and features of the terascale DNS code available.

To develop the first complete, time analysis of high scalar dissipation rate regions, we visualize these regions in 3D and allow the scientists to change the region selection parameter in real time. Figure 16 shows the stoichiometric mixture fraction surface superimposed with the selected high scalar dissipation rate regions.

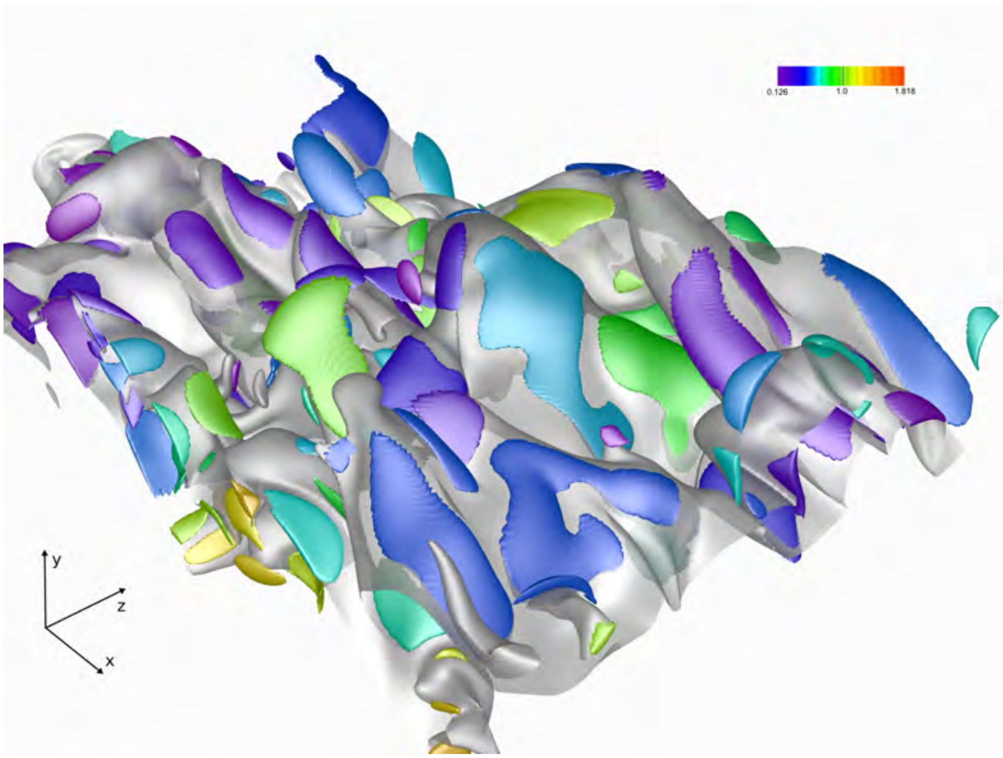



Figure 16: This image shows the stoichiometric mixture fraction surface superimposed with the selected high scalar dissipation rate regions.

### Correspondence Between Features and Combustion

- Compare quantities averaged over the portion of the stoichiometric mixture fraction surface inside and outside the features, normalized by the average over the entire surface.
- Comparison for all features intersecting the surface shows only small differences; average over portion of surface outside the features closely approximates the average over the entire surface.
- Significant change in conditional reaction with conditional scalar dissipation within feature suggests that not all features have high enough scalar dissipation to experience extinction.
- Comparison including only those features with large conditional chi shows much larger effect.

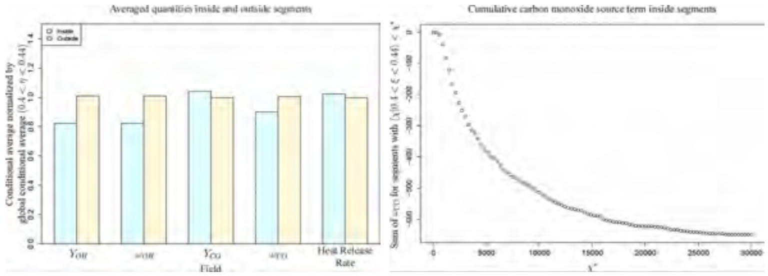



Figure 17: Comparison for all features intersecting the surface shows only small differences; average over portion of surface outside the features closely approximates the average over the entire surface.

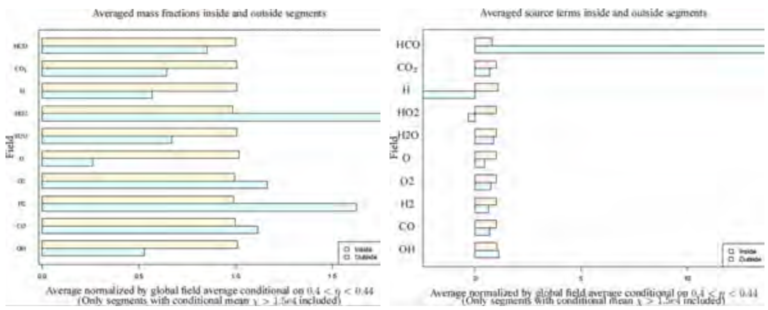



Figure 18: Comparison including only those features with large conditional chi shows much larger effect.

**Temporal History**

- Connecting features with spatial/temporal overlap allows time tracking.
- Following the ten largest features at 20 jet times (vertical line) forward and backward in time produces the "tracking graph" (Figure 19).
- These features are the result of merges earlier in time.
- Later in time, the selected features split.
- Prior to 20 jet times, some of these features have been part of interacting families.
- The scalar dissipation rate magnitude for the tracking families fluctuates significantly.
- The fluctuations have a longer timescale for the families from this sample which have experienced the largest scalar dissipation rates (Figure 20).

**Alignment between strain field and scalar gradient For non-reacting flows** The scalar gradient is preferentially aligned with the most compressive turbulent strain. Within the features, the preferential alignment is more pronounced than outside (Figure 21).

**Impact** Scientists were able for the first time to use a full 3D segmentation of extinction regions and follow in time their individual evolution, with detailed creation, destruction, merge and split events. This will allow a better understanding of complex dynamics of turbulent combustion processes and ultimately allow the design of cleaner and more efficient engines.

At the SciDAC 2008 program meeting, Jackie Chen's invited presentation included discussion that she was now seeing new features in her data never before possible. It is clear that this work is providing the ability for new scientific knowledge discovery.

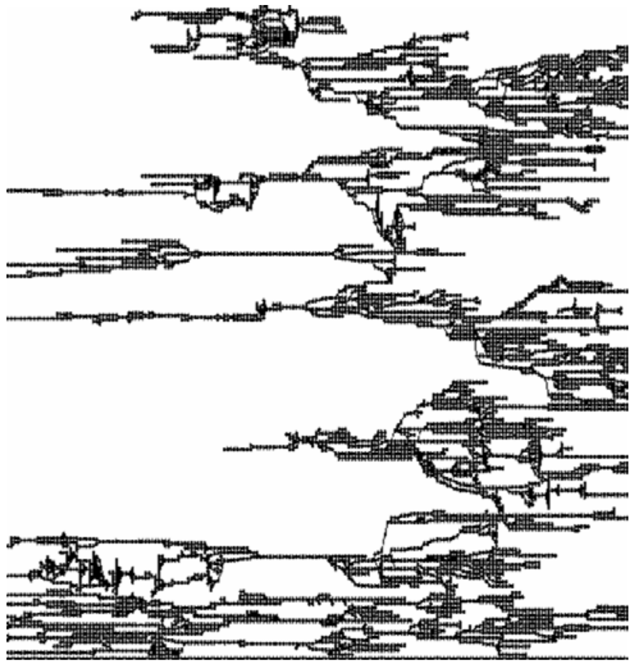


Figure 19: Connecting features with spatial/temporal overlap allows time tracking. Here, connected features are shown in graph form.

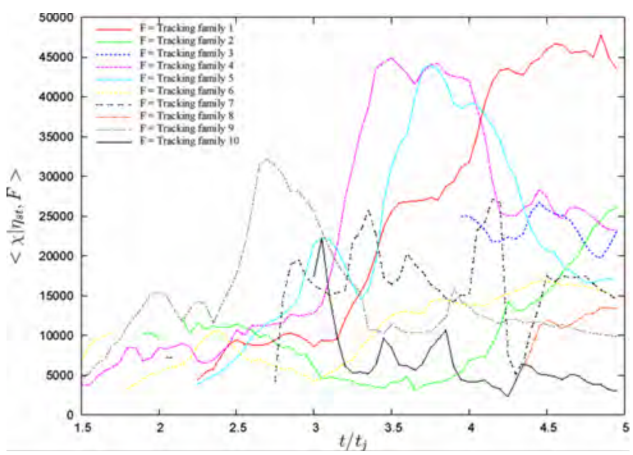



Figure 20: Scalar dissipation for tracked features.

## Future Work

Continue collaboration with stakeholder to prepare paper analyzing the combustion aspects of the research.

> These methods have enabled us, for the first time, to robustly track over time large numbers of extinction or ignition regions, which is essential to the development of a better understanding of the flame extinction/reignition and stabilization dynamics. – *Jackie Chen, SNL-CA.*

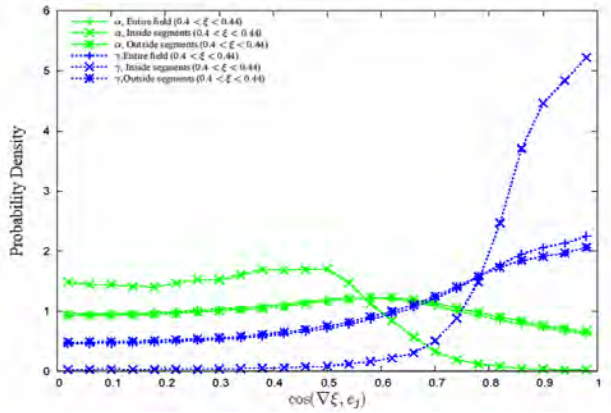



Figure 21: For non-reacting flows, the scalar gradient is preferentially aligned with the most compressive turbulent strain. Within the features, the preferential alignment is more pronounced than outside.

### 3.9 Fusion: Framework Application for Core-Edge Transport Simulations (FACETS)

The Framework Application for Core-Edge Transport Simulations (FACETS) SciDAC project is a multi-physics, parallel framework application supporting various fusion efforts. VACET has been an enabling factor in a number of fronts for FACETS visualization and analysis needs.

One of the R&D efforts by the VACET team has been focused on analysis tools which are critical for the kind of data the FACETS team generates. On that front, enhancements have added advanced, high-order, accurate integrators, better user controls, and parallel, scalable algorithms for calculating streamlines in petascale environments. These are also enabling integration of new Poincaré plot and operator components for scientifically meaningful analysis of their data.

As they retire old visualization programs in use by the FACETS community, a major desire is to have the ability to create domain-specific tools with a simpler, more targeted user interface familiar to the community. By re-architecting portions of VisIt, it is now possible to write applications that contain all the power and advanced features of VisIt, but packaged with a new, fully customized interface. A demonstration of this is seen in figure 22, where two of VisIt's many plots have been wrapped in a standalone application with simple controls.

Of course, VisIt must be able to read the data of the FACETS framework before it can be plotted or analyzed. Through its plug-in mechanisms, VisIt has enabled the FACETS project to write custom file format readers, and has added features to enable those plug-ins to have customizable options designated by the user at runtime. And close collaboration with the VACET team members has also aided parallelization of their VizSchema file reader plug-in, which is now routinely used to read large data files of 100GB or more.

> I am sold on VisIt and the VACET team. They are providing a great service to our community in not only providing a tool, but also in providing usable tool [sic] that helps scientist [sic] extract knowledge from the huge amounts of data that are being generated by SciDAC applications now in use. – *John Cary, Tech-X Corporation*

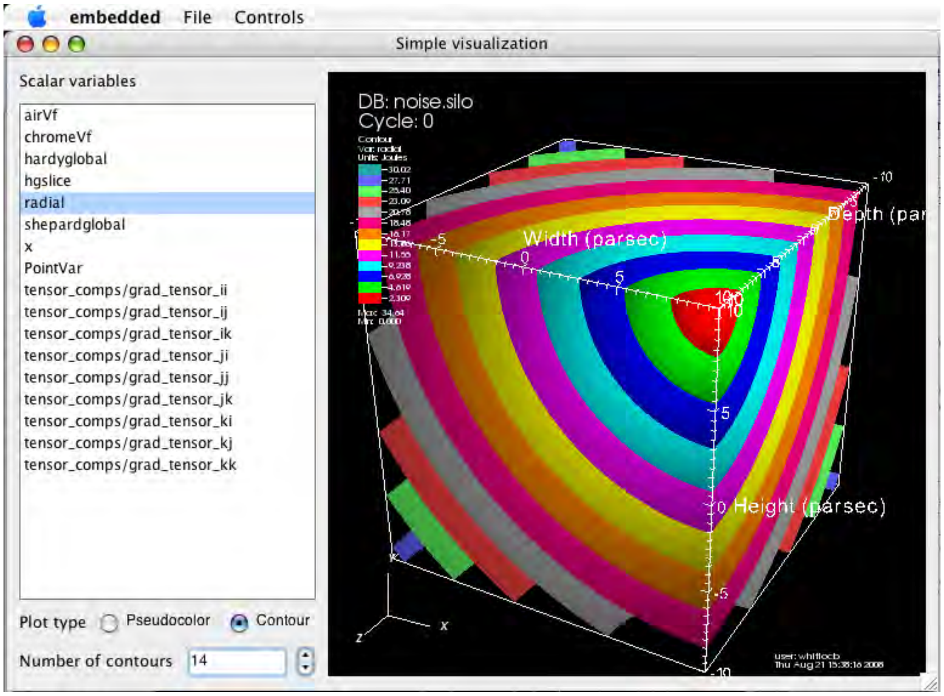



Figure 22: Screen shot of a VisIt "skin," an interface customized for individual customer needs.

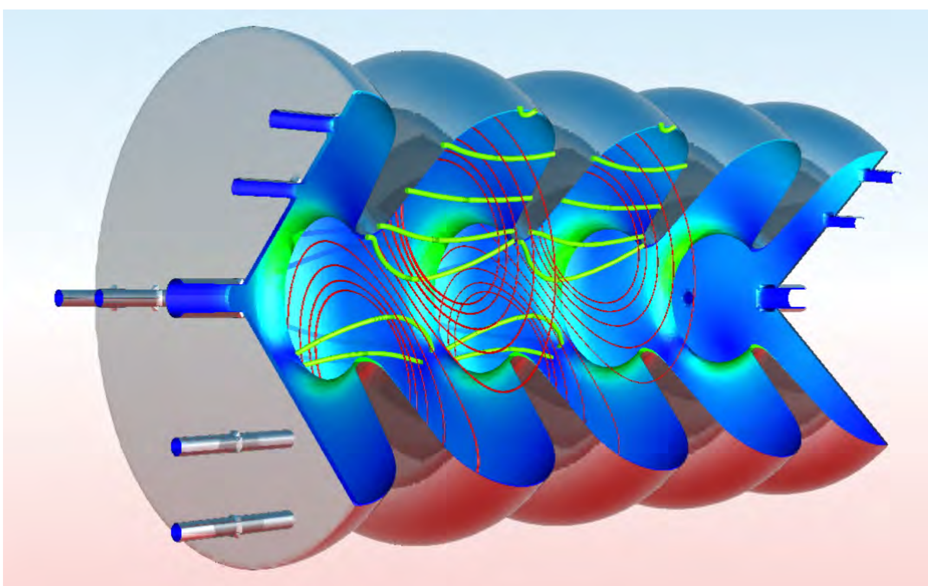


Figure 23: One frame of a movie that inspects a crab cavity from a VORPAL simulation. This movie received a People's Choice Award at SciDAC 2008's "Vis Night." It was generated solely by FACETS group members from Tech-X, using VisIt.

FACETS has been adopting VisIt as their tool of choice for these tasks, and as expected, the VACET team has responded with many forms of user support and bug fixes. Some examples include:

- Adding secure tunneling to VisIt's networking infrastructure to support FACETS environments.
- Support for running on large machines at NERSC
- Adding capabilities for limiting user exposure to extraneous database formats
- Simplifying visualization session persistence and recovery
- Continuous email support for both usage and development queries
- Assistance with image generation and analysis
- Porting to new FACETS machines and operating systems as they are deployed
- Site visits and outreach to Tech-X headquarters
- VACET representation at FACETS All-Hands Meetings

Due to these efforts from VACET, the PI for the FACETS project, John Cary, says he is "sold on VisIt." With VisIt they were able to produce an award-winning movie, independent of any specific assistance from VACET. Figure 23 shows a frame from this movie, which won an OASCR People's Choice Award at the 2008 SciDAC meeting.

### 3.10 Fusion: Particle Path Analysis and Visualization

Several SciDAC fusion projects use particle-in-cell codes to model scientific phenomena. VACET is working with several of these teams by focusing on the question "what are the visual data analysis needs of fusion researchers who use particle-based codes to study fusion?" Our work in this area combines visualization R&D with software engineering to produce effective, production-quality visual data analysis software infrastructure aimed at meeting the needs of this broad cross-section of the fusion community.

To achieve these objectives, our team is working with fusion researchers directly in a VACET-stakeholder relationship, as well as through a recently awarded SciDAC Science Application Partnership that is led by Allen Sanderson (Utah) from VACET and representatives from two different SciDAC fusion projects.

The fusion stakeholders are currently running simulations that use and generate millions to billions of particles, with each particle containing multiple scalar and vector data (multivariate data). They would like to have the ability to explore the nature of the particle orbits in an interactive manner. It is impractical to view billions, much less millions, of particles at one time and glean any insight. As such, the physicists would like to have tools that allow them cull the particles using a user defined query or other statistical tools.

> Through a productive collaborative effort with our science team, the VACET Center has had a positive impact on our project. The initial prototype tools developed helped us better understand some of the transport phenomena linked to the nonlinear interactions between charged particles and waves in the plasma. . . . We are now looking forward to the full deployment of these tools later this spring, especially now that we are about to commence simulations that will contain in excess of 20 billion particles, which will be very difficult to analyze without the VACET collaboration. – *Stephane Ethier, PPPL.*

At the same time as the particles are displayed, physicists are interested in seeing the particles in context of data that is not associated with the particle but is part of the simulation. This data may be scalar (electric potentials) or vector data (magnetic field) and may have its own visualization requirements. For instance, the scalar data may be viewed using a variety of techniques from volume rendering to slicing.

It turns out that selection/subsetting and subsequent visual data analysis of particle based data is nearly identical for fusion as accelerator applications (see Section 3.1). Therefore, we are able to leverage an existing investment in infrastructure and adapt it for use in another application area.

Major accomplishments over the course of this project are:

- Ascertain "particle data analysis" needs common across multiple fusion projects, formulate work plans from these that will result in new capabilities that meet the needs of a large cross-section of the fusion research community.
- Joint funding with fusion researchers in the form of a SciDAC Science Application Partnership to augment existing funding to support the work.
- Design and implement a data model and parallel I/O strategy suitable for use with petascale-class datasets and visual data analysis software infrastructure.
- Generate a working prototype (Figure 24) that supports interactive multivariate subset selection and visualization in physical space.

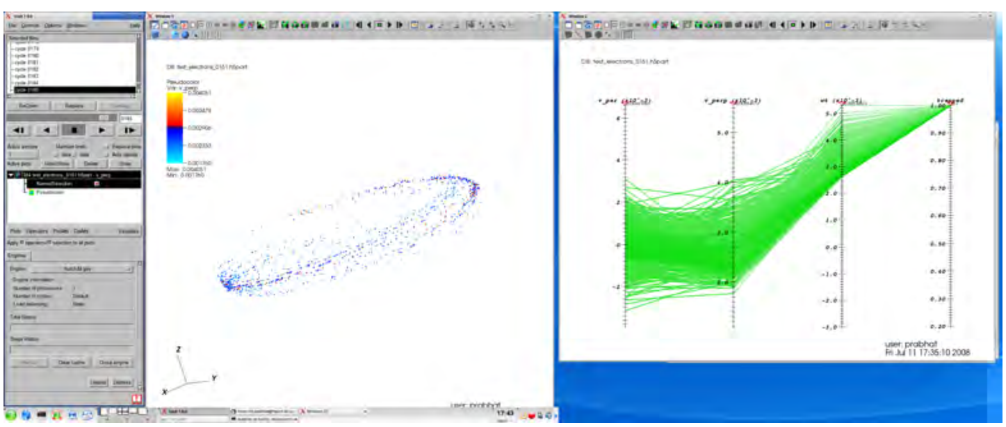


Figure 24: Screen snap shot of query based fusion particle visualization using parallel coordinates.

### 3.11 Fusion: Magnetic Field Analysis

Here, VACET is using the same approach as for the fusion community that uses particle-based models to study phenomena: we have coordinated efforts with several fusion teams that are interested in studying and analyzing the magnetic fields that form in burning plasmas. The objective is to focus on a handful of needs common across teams, then deliver capabilities that meet those needs in production-quality form.

Physicists are currently studying the effects of magnetic islands that form in plasma. These islands cause defects in the magnetic field and the current flow resulting in contact between previously separate regions. This contact results in "hot areas" coming into contact with "cool areas,"

which leads to core cooling. Physicists would like to have tools that allow them be able to automatically generate Poincaré maps of the magnetic field and detect the island formation and track them over time.

> The VACET Center has had a positive impact on our project by supporting our team as we move towards utilizing VisIt as our primary visualization and analysis tool. ...The quality of the initial analysis tool has generated interest no only within our group but within other fusion SciDAC's as well that are headed by Don Batchelor and Bill Nevins. ...VACET is a valuable asset to the scientific community and we look forward to our continued collaboration.
> *– Stephen C. Jardin, PPPL*

At the same time as the Poincaré maps are displayed, physicists are interested in seeing them in context of data that may or may not be associated with it, but is part of the simulation. This data may be scalar (electric potentials) or vector data (magnetic fieldlines) and may have its own visualization requirements. For instance, the scalar data may be viewed using a variety of techniques from volume rendering to slicing.

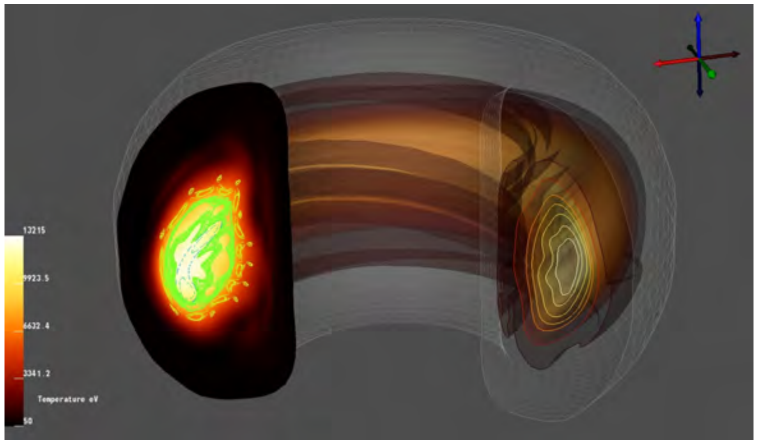
Major accomplishments over the course of this project are:

- Ascertain "magnetic field data analysis" needs common across multiple fusion projects, formulate work plans from these that will result in new capabilities that meet the needs of a large cross-section of the fusion research community.
- Joint funding with fusion researchers in the form of a SciDAC Science Application Partnership to augment existing funding to support the work.
- Generate a "Poincaré plot library," port to VisIt and demonstrate in other standalone tools (e.g., gnuplot).
- Leverage related work in VACET streamlines research (Section 4.1) to take advantage of alternate integration techniques and parallel computing infrastructure.
- Extensive outreach: tutorials (e.g., SciDAC 2007, 2008 meetings) and workshops (e.g., at PPPL for fusion researchers in Fall 2008) to help this community migrate to modern, production-quality visual data analysis software infrastructure.

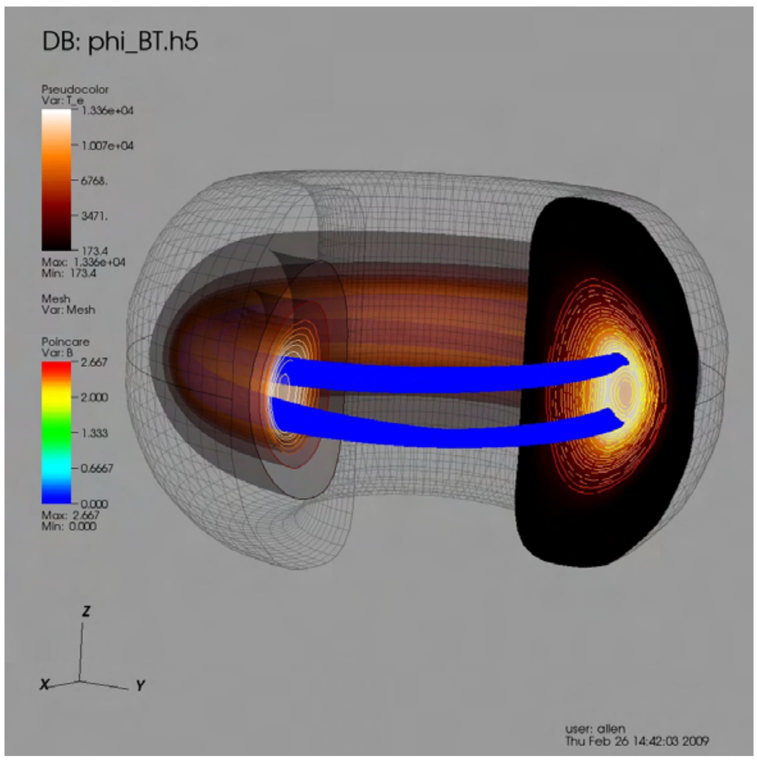
### 3.12 Mathematics: Applied Partial Differential Equations Center

The SciDAC Applied Partial Differential Equations Center (APDEC) is developing a collection of algorithmic and software components that can be assembled into simulations for modeling complex multicomponent physical systems. Their approach relies on solutions to partial differential equations, and they rely extensively on Adaptive Mesh Refinement (AMR) grids. AMR offers several unique advantages, the primary being the ability to efficiently and adaptively support computations over vast ranges of spatial and temporal scales across several orders of magnitude.

APDEC partners with several science applications to help assemble simulation codes that use AMR and PDE techniques to model physical phenomena: combustion, accelerator, and MHD tokomaks (Figure 26. Our approach is to partner with APDEC to deliver the world's best AMR visual data analysis software infrastructure and thereby provide a positive impact to all of APDEC's stakeholders now and in the future.




(a) Here we see the breakup of the magnetic field into a series of island chains (left half) along with isosurfaces of the plasma temperature (right half).




(b) Screen shot of a Poincaré plot of magnetic fusion data created with VisIt.

Figure 25: Visual data exploration of magnetic fields, including visual data analysis to detect and display the "breakup" of the magnetic field.

To that end, VACET made significant enhancements to VisIt to enable the Applied Partial Differential Equations Center (APDEC) to migrate from ChomboVis, their homegrown visualization tool, to VisIt. By using VisIt, APDEC can focus their efforts exclusively on their center's mission, as well as gaining more functionality, including, most importantly, parallel visualization algorithms that will allow them to view their biggest simulations. Application specific visualization tools (like ChomboVis) typically have a streamlined interface and custom capability, which makes migrations

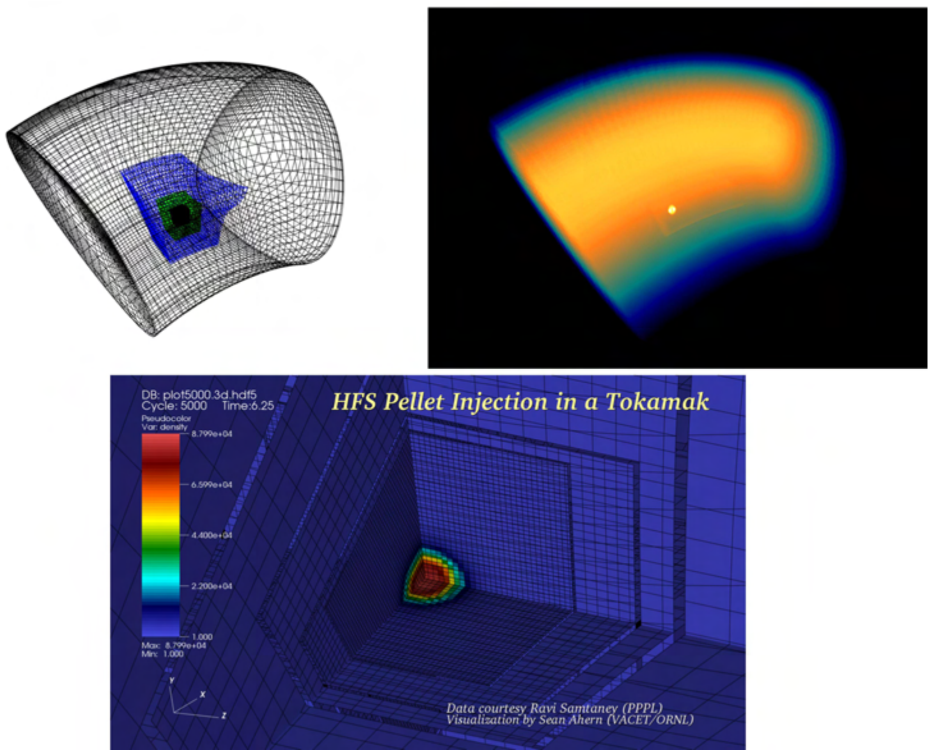



Figure 26: Pellet injection is a mechanism for the refueling of burning plasma reactors (e.g., ITER). The physical processes are inherently multiscale which has prevented effective simulations in the past. Block-structured AMR allows researchers to efficiently simulate pellet injection and investigate in detail the fueling and mass redistribution processes in a tokamak. Top Left: Block-structured adaptive curvilinear meshes to resolve the pellet. Top Right: Volume rendering of pressure shows hot spot due to electron heating (visualization: H-W Shen & S. Martin, Ohio State University). Bottom: Density slices show effective resolution of the pellet (visualization: S. Ahern, ORNL/VACET).

difficult. To ensure the migration was a success, VACET had to significantly enhance VisIt's feature set for Adaptive Mesh Refinement (AMR) grids and also provide more customizability to VisIt's user interface, to allow APDEC users to have an equally streamlined process.

Although much of the effort required to transition APDEC to VisIt would not qualify as "publishable research," we felt it was critical to dedicate resources to have a positive impact on this scientific community. Further, these changes positively affected VisIt on the whole and therefore benefited the entire VisIt community. Finally, by making this initial investment, we have laid the groundwork needed to be able to quickly deliver new capabilities emerging from our R&D pipeline to APDEC in a production-quality, parallel capable form.

Specifically, we:

- We optimized VisIt's rendering algorithms, gaining a factor of five in both rendering time and memory footprint on AMR meshes.
- We significant tuned the memory footprint and performance of VisIt on large number of patches.
- We added a spreadsheet capability that allow for Excel-style viewing of AMR patches.
- We integrated VisIt with APDEC's debugging cycle, so that APDEC developers can invoke

VisIt from the "gdb" debugger and visualize bad patches.

- We added "macro" capabilities and "Python callbacks" so that they could easily customize their environment and perform actions in a single click.
- We added capabilities for AMR grids with irregular meshes ("mapped grids"), which was important for accurate viewing of their data.
- And we have responded to over 30 requests for minor interface changes, enhancements, and bugs.

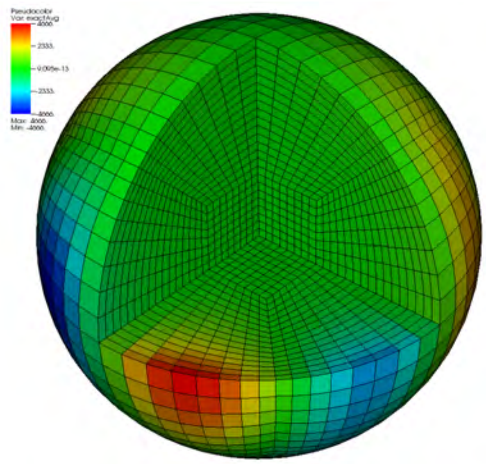


Figure 27: Data from APDEC rendered correctly as a "mapped grid."

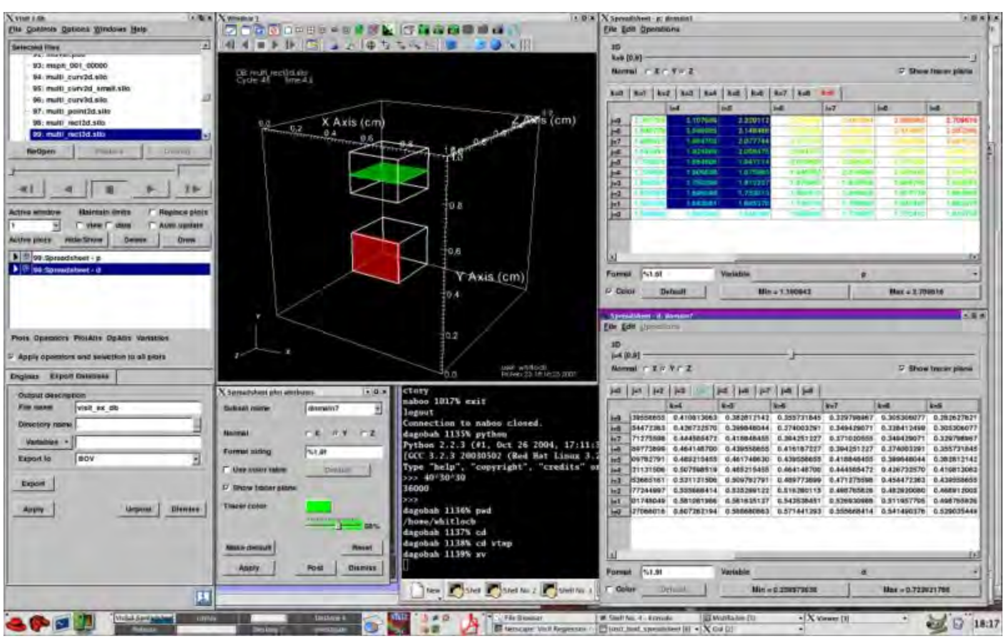


Figure 28: Excel-style spreadsheets were in an important usability feature to enable the migration of ChomboVis users to VisIt.

# 4 Field-Leading Research

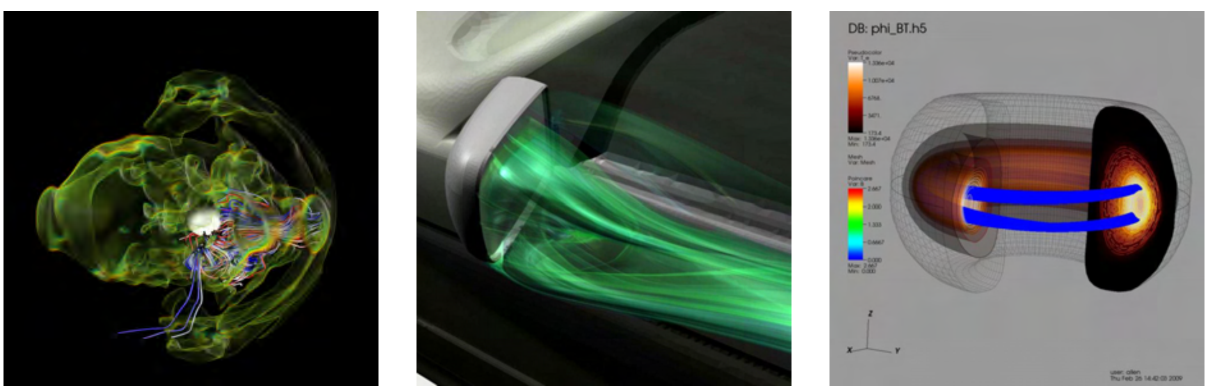
A portion of VACET's technical portfolio focuses on research activities, both applied and basic. Our research agenda is driven primarily by science stakeholder needs – all research projects in our portfolio can be traced to a specific science stakeholder problem. This section provides an overview of some of our research efforts.

## 4.1 Streamlines and Stream Surfaces

A long-standing gap in visualization technology in general is the computation of streamlines[7] in parallel and on multi-grid domains. The objective for this element of VACET's research portfolio is to find a general solution to this problem that is suitable for use in solving several different science stakeholder needs and in a way that is suitable for deployment in a variety of different software applications. The target is a parallel capable "streamlines engine" that meets multiple stakeholder needs and that advances the field of high performance visualization.

Our new parallel-capable, multi-grid aware streamlines engine has been deployed in VisIt (version 1.10), VACET's production-quality parallel capable visual data analysis software infrastructure. This important new capability will help scientists gain deeper understanding into complex, time-varying and multi-grid vector field data produced by large-scale simulations.

The "streamlines engine" that results from this R&D effort has many potential uses under the rubric of visual data analysis. First, it can compute streamlines and advect particles through flow fields. Second, it can compute stream surfaces, which are an easier-to-comprehend mechanism for conveying structure in 3D flow fields. Third, it can serve as the basis for tools that need to compute flow paths through flow fields, tools like the Poincaré analysis common in fusion sciences for studying the quality of magnetic fields in burning plasmas.




(a) Streamlines through supernova collapse simulation data. This image appeared on the cover of the SciDAC Review, Special Issue, 2009.

(b) Stream surface showing 3D air flow structure around a car mirror.

(c) Poincaré plot computed inside VisIt using the new streamlines engine.

Figure 29: The new VACET streamlines engine computes stream lines and surfaces on multi-grid domains and in parallel, but also serves as the basis for more advanced visual data analysis tools required by our science stakeholders.

This research, development, and deployment work was conducted by VACET researchers from

[7] A "streamline" is a curve that is tangent to a flow field along its length.

four different institutions (UC Davis, ORNL, LBNL, and LLNL). See Section 8 for a highlight of this work.

## 4.2 Embedded Boundary/Material Interface Analysis

In many applications it is necessary to reconstruct or track the boundary surfaces, or "interfaces," between multiple materials that commonly result from multi-fluid calculations. This problem, where the generated data sets have the characteristic that each cell contains a "volume fraction" representing the percentage of each material contained in the cell, now arises in a variety of applications, and is frequently called the embedded boundary problem. Having support for embedded boundary reconstruction and display is one of the top priorities for one of our science stakeholders, the SciDAC Applied Partial Differential Equations Center. Its solution is a long-term research issue for VACET.

The challenge is to utilize the material fractions in each cell to reconstruct the boundaries between materials. VACET researchers have been developing a new approach to this problem. Before tackling the problem, we consulted with material interface reconstruction experts from Lawrence Berkeley, Lawrence Livermore, and Los Alamos. We strove to design a broad solution that could affect many stakeholders. Various laboratories require different solutions to this problem: some require a solver that takes only general volume fractions as input and outputs material boundaries. Another requires two-material solutions on AMR meshes. All require parallel implementations and the ability to work on diverse grids.

We have developed a new solver based upon active contours. Active contours are piecewise linear curves that move toward a specific target boundary. They work by balancing three forces, a force that attempts to keep the points of the curve equidistant, a force that attempts to limit the curvature of the curve, and a volume aware force that adjusts the curve motion according to volume constraints. Once an initial approximation of the curve is generated, the active contour quickly converges to the boundary. The model works on two- and three-dimensional data sets. Initial results have proven to be very accurate, with error levels for these models several orders of magnitude lower than previous methods. Example results are shown in Figure 30 comparing the results of a legacy algorithm with our new implementation.

Material interface reconstruction algorithms are judged by many metrics: accuracy (in terms of preserving volume fractions), smoothness of output, quality of topology, size of output (in terms of number of primitives), and, of course, execution time. We designed our algorithm to perform well with these metrics. However, we plan to demonstrate its effectiveness before moving forward. If it is successful, we will extend the algorithm for widespread use by a broad community: adapting it for use on varying grid types and extending it for use on multiple materials. After that we will productize the algorithm and deploy it within VisIt.

## 4.3 Hardware Accelerated and High Quality Volume Rendering

### 4.3.1 SCIRun Library for Volume Rendering (SLIVR)

Although we have targeted VisIt as the production visualization tool for many of our partners, the image quality produced by VisIt's volume rendering algorithm has proven to less than desired for many of our projects. Therefore, we undertook a Center-wide project aimed at "technology transfer" of an existing implementation of a high-quality volume rendering engine into VisIt.

The approach was to first convert SLIVR: the SCIRun Library for Interactive Volume Render-

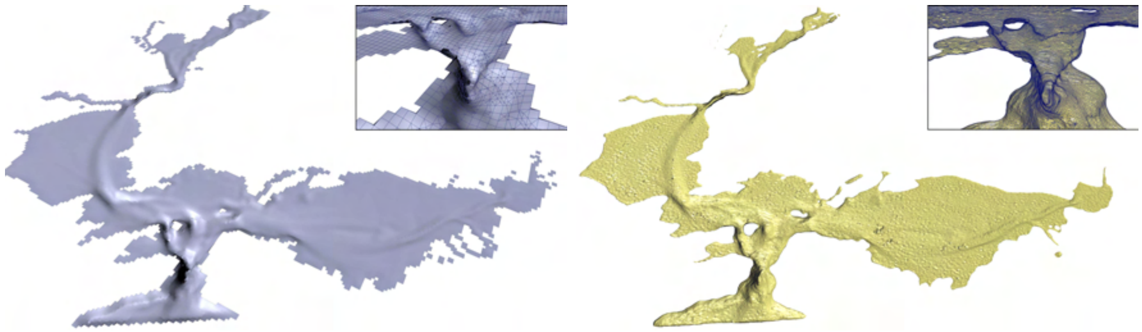
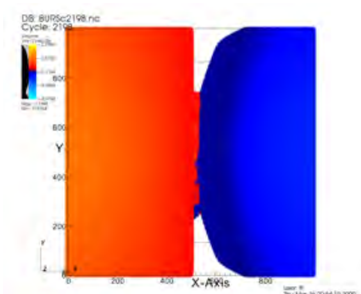


Figure 30: Comparison of old (left) and new (right) algorithms for performing material interface reconstruction.

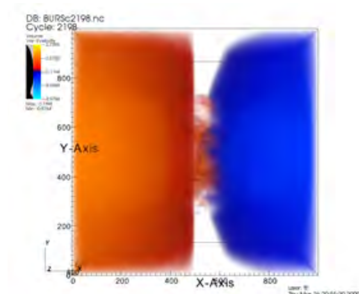
ing[8], into "library form" for use by other applications, then add this capability as a new volume rendering option into VisIt. This project, which involves VACET personnel from Utah, LLNL, and ORNL, is a successful example of focusing software engineering effort to bring a successful research prototype into production use.

SLIVR is now completely integrated into VisIt, and are available to the worldwide scientific community as of VisIt version 1.11.

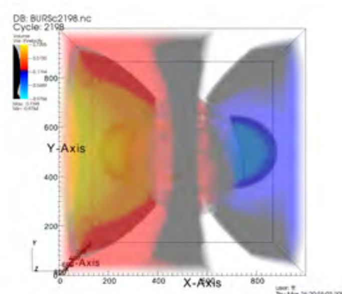
Figure 31 shows a comparison of three volume rendering algorithms in VisIt applied to astrophysical simulation data.




(a) Splat-based volume rendering.




(b) 3D texture-based volume rendering.




(c) SLIVR volume rendering.

Figure 31: Comparison of three different volume rendering algorithms in VisIt: splat-based (a), 3d texturing (b), and SLIVR (c). The source data, provided by J. Blondin (N.C. State), is from a simulation that models the core collapse of a supernova, the source of the diversity of elements in the universe. These images show the X-component of material velocity around the expanding supernova shockwave. The SLIVR renderer reveals much more detail than the other two algorithms.

#### 4.3.2 Tuvok: GPU-Accelerated, Large-Scale and High Quality Volume Rendering with Novel User Interface/Interactions

Scientists from the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) have developed cutting-edge techniques for visualizing large volume data. Through an emphasis on real-time performance, scientists are encouraged to explore their data and develop new theories.

[8] See `http://slivr.sci.utah.edu`.

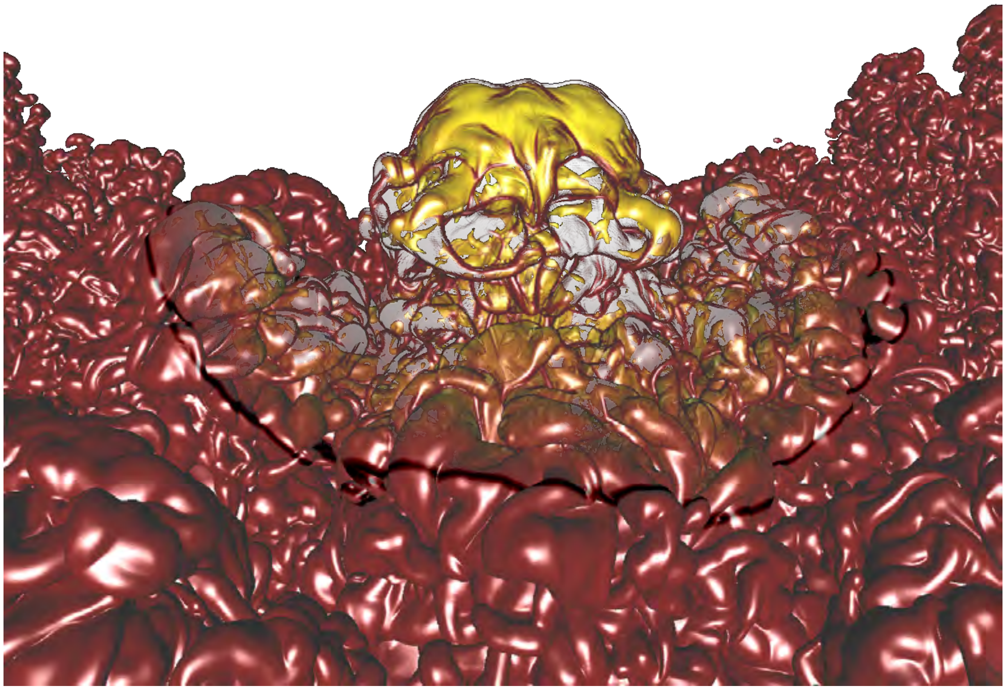


Figure 32: Tuvok using its "ClearView" rendering mode. ClearView suits the focus-and-context exploration model, common in large-scale scientific visualization, through an intuitive model of a lens which can see 'in' to the dataset. In this data, scientists can rapidly discern the internal behavior of a Richtmyer-Meshkov simulation at the interface, while simultaneously viewing the global structure of the fluids.

To achieve this level of performance, the Tuvok volume rendering library leverages the immense computational resources available on recent graphics processing units, and will be paired with more traditional parallel computing resources. Ongoing work is underway to integrate these tools into production-quality general-purpose visualization and analysis tools, such as the VisIt software package.

A common way to gain insight into volumetric data is through a technique dubbed "volume rendering". Due to its ability to efficiently communicate three dimensional structures, volume rendering is applicable to a wide variety of disciplines.

Due to its inherently three dimensional nature, a volume renderer must examine a large amount of data before displaying an image. As simulation data grow in size, traditional volume rendering applications are strained under the increased workload. To combat this issue, scientists from VACET have developed Tuvok, a volume rendering library that harnesses the immense computational capabilities of commodity graphics processing units (GPUs). Using Tuvok, scientists can use a typical workstation rather than a small supercomputer to visualize the 8 gigabyte Richtmyer-Meshkov dataset (Figure 32.) With Tuvok's level-of-detail features enabled, the dataset can be interacted with in real time, at rates exceeding 60hz.

Future work will include: (1) porting Tuvok into VisIt for use as another volume rendering method; (2) exploring methods for exploiting multi-GPU systems (e.g., GPU clusters) to reach

higher levels of performance and on larger datasets; (3) leverage VisIt's internal data management infrastructure to gain better traction on large-data visualization problems. In the long run, Tuvok will replace SLIVR as VisIt's high quality volume rendering algorithm.

## 4.4 Topological Analysis – Feature Detection, Tracking, and Analysis

Two of our most significant research contributions in this area are detailed in Sections 3.7 and 3.8. The basic idea is to study features in data with the aim of gaining better understanding of relationships between input parameters and their impact on models of physical and/or chemical systems. This type of understanding is simply not possible with traditional visualization algorithms that directly produce images from scientific data.

## 4.5 Query-Driven Visual Data Exploration and Analysis

The term "Query-Driven Visualization" (QDV) means focusing visualization and analysis processing, along with subsequent rendering and human cognition, to the subset of data that is "scientifically interesting." Stated differently, while it is possible to make an image of a TB or PB sized dataset, the resulting image will be of little value due to its spatial and depth complexity. Instead, QDV is akin to "finding and showing needles in haystacks." The technical challenges are that the haystacks are growing larger and more numerous, and it is growing increasingly difficult to describe the needle as scientific data grows more complex.

VACET researchers introduced this field-leading concept in 2005 (before VACET started), and have been continuing R&D in this area to achieve VACET and SciDAC mission objectives.

**Accomplishments**

- Coupling high performance index/query infrastructure[9] with advanced user interface technology, VACET researchers were able to provide a new capability to the accelerator research community: the ability to quickly find, track, and analyze particles undergoing wakefield acceleration in multi-TB datasets. This new capability runs in seconds or minutes, compared to hours or days required by a previous process. See the two-page article in Section 8.
- Working researchers from the Institute for Ultrascale Visualization and the SciDAC SDM Center, we developed a new indexing technique that runs in parallel on a GPU. This new capability accelerates (reduces) multivariate search time.
- We developed a novel method for index/search and subsequent visual data exploration of AMR-based data.

## 4.6 Using Machine Learning to Identify Beam Particles

The scientific stakeholders on this effort[10] are both experimentalists and computational scientists. The primary objective for this project is to devise a mechanism for automatically finding particles that are undergoing acceleration in a laser-wakefield simulation dataset. This capability would help to automate a process that is currently performed manually. In the long run, this capability will certainly help to accelerate data understanding. As it evolves, this technique will likely be

[9]FastBit from the SciDAC Scientific Data Management Center.

[10]C. Geddes and E. Cormier-Michel (LBNL), P. Messmer (Tech-X) are INCITE awardees at NERSC and part of the SciDAC COMPASS project.

integrated with other tools (e.g., high performance visual data analysis) as part of a broad set of HPC tools for scientific knowledge discovery in accelerator science.

**Accomplishments**

We devised a methodology for automatically identifying beam particles, ie., those undergoing wake-field acceleration, from data produced simulation:

- The methodology identifies "bunches" of particles likely to be part of the beam as those having both high momentum and are in close spatial proximity.
- For each bunch of particles from all timesteps' worth of data, we use a graph algorithm to track bunch movement and evolution across timesteps.
- Separately, we use fuzzy clustering to classify particles as either "beam" or "non-beam," where classification is continuous rather than discrete.
- Results of fuzzy clustering are compared with those from the space/momentum classification stage. Where there is agreement between these two models, we have a high quality beam.

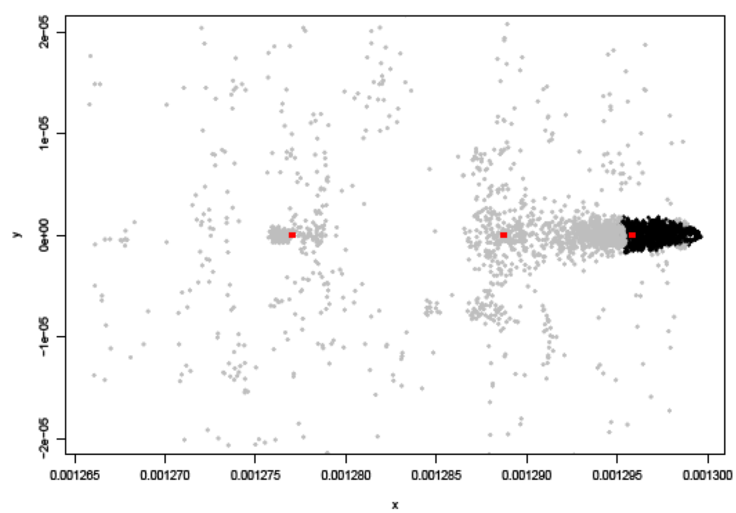


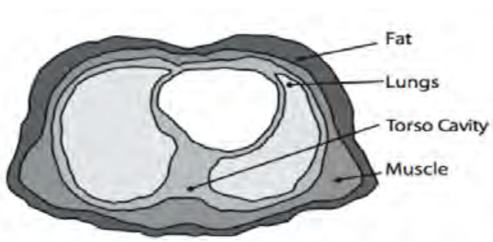

Figure 33: Beam-point candidates combined with fuzzy clustering. Beam points, which are comprised of high-momentum particles having spatial coherence are shown in red. The results of fuzzy clustering shows beam particle candidates in black, and non-beam particles in gray.

## 4.7 Uncertainty Visualization

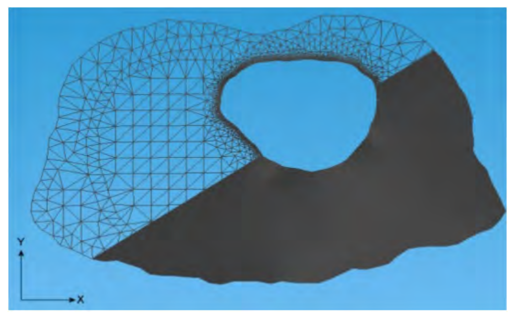
Uncertainty information is an important characteristic associated with much of the data scientists encounter. Such information is typically included with the data as 2D charts and graphs, however incorporating uncertainty into visualization techniques has proved quite challenging. Recent work examines a class of uncertainty data that is characterized as a set of probability density functions (PDFs) defined across a triangular mesh, and explores ways at visually presenting this data.

The data used in this work comes from the sensitivity analysis electrical conductivity within computational models of bioelectric fields of the heart. Such an approach quantifies the sensitivity of the electrocardiographic forward problem by creating a mathematical model to reconstruct a biological experiment in which the voltages on the human torso are estimated based on the input electrical conductivities (from previous work). The simulation stochastically varies the input conductivities of different tissues such as fat, lungs, or muscle and examines the resulting changes in potential across the torso.

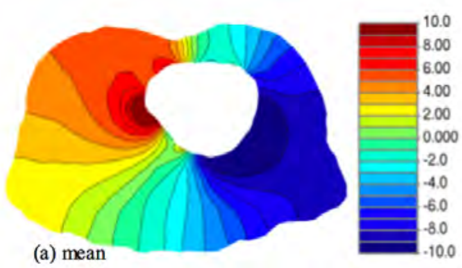
Our recent visualizations strive to investigate the complex sensitivity analysis data. The data is defined on the classified torso mesh (shown below in Figure 34). The mean and standard deviation of the data are shown colormapped onto the torso data space, and the first approach we explore is to combine the mean and standard deviation into a single visualization. We achieve this by simply encoding the mean into a heightmap, and colormapping variance onto this map. This provides for a simple way to show two variables of the dataset, and quickly find locations of large and small average values of the data, as well as high variance values.



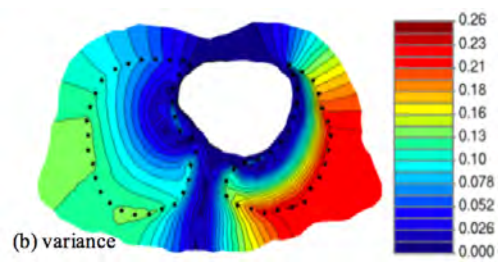

(a) The classified human torso.




(b) The torso mesh space.



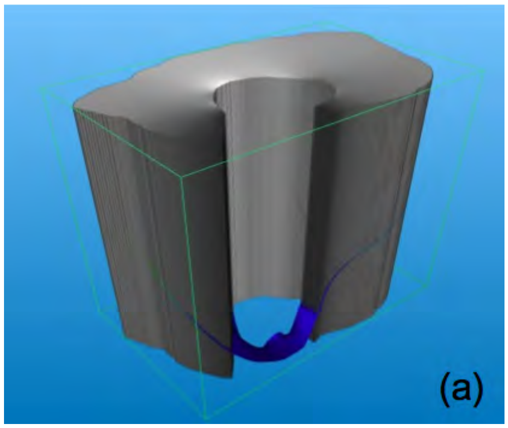

(c) Mean colormapped onto the torso.



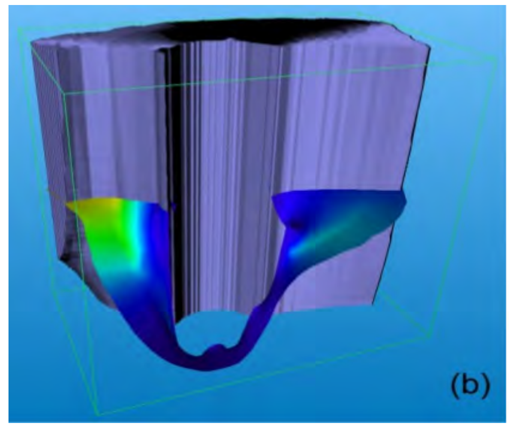

(d) Standard deviation colormapped onto the torso.

Figure 34: Comparison of legacy (left) and new (right) techniques for computing embedded boundaries.

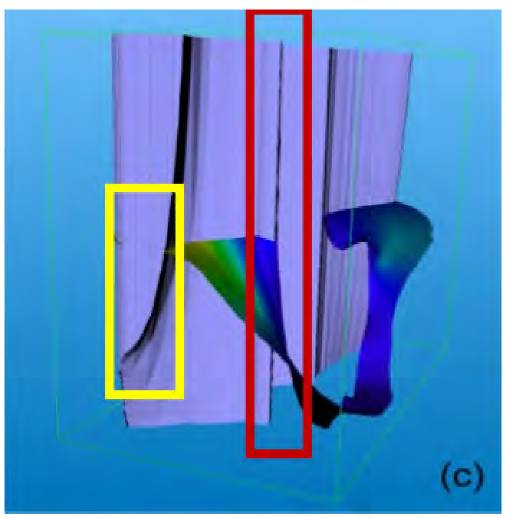
The next approach we investigate is to create a volume by stacking the output potentials for every input conductivity. This volume can be volume rendered, or iso-surfaced. The structure of the iso-surface is of most interest. An iso-surface that falls straight down indicates the potential at this point in 2D space does not vary when changing; thus, it is considered independent at that point. A bending iso-surface indicates areas of high dependence on the input conductivity. We can also see that even in the region with the highest variance only small conductivities result in potential changes. This relationship becomes visible using our new visualization technique (Figure 35).



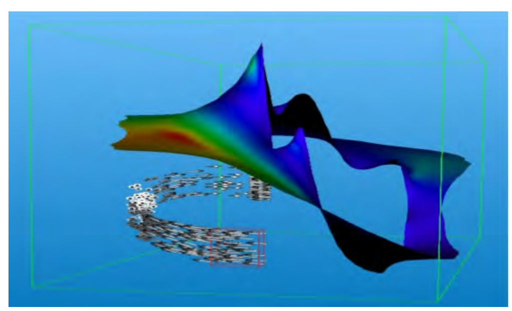

(a) Direct volume rendering of the potentials volume superimposed onto the mean height-field.




(b) Isosurfacing the potentials volume.

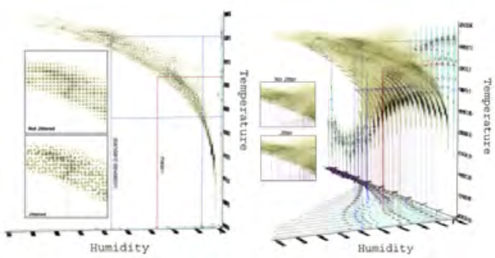



(c) Iso-surfacing of the potentials data. Bends in the iso-surface (yellow box) show areas with high dependence on the input conductivity.

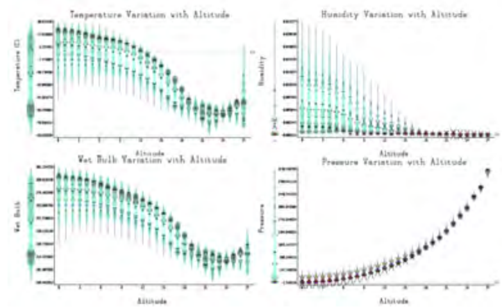


(d) Particle tracing on the potentials volume.

Figure 35: These images show the results of recent research in the field of uncertainty quantification and visualization.




(a) Joint mean (red lines), joint standard deviation (blues lines) and the joint histogram for a single (left) and multiple (right) categorical datasets.




(b) Temperature, humidity, wet bulb, and pressure data taken at a single moment in time across the world. Data courtesy of the Canada Meteorological Centre.

Figure 36: Visualizing statistics using different new techniques.

**Visualizing Summary Statistics and Uncertainty**

An important visualization research problem is to effectively convey uncertainty information along with traditional visual data representations. Recent work investigates the problem from a graphical data analysis standpoint. By using descriptive statistics to summarize both characteristic features of a data distribution and measures of uncertainty, we can achieve a more cohesive understanding of the information. In this work, we reexamine the box plot and its relatives and develop a new hybrid summary plot that combines moment, cumulant, and density information along with higher order descriptors that rely on distribution fitting. In view of the important role summarizing plots has in decision making, our work focuses on using advanced visualization techniques to incorporate additional descriptive parameters, while simultaneously improving the comprehensibility of summary plots.

The approach this work takes to understand and visualize uncertainty data focuses on methods for quantitatively displaying the underlying statistics which describe the uncertainty of a dataset. The measures typically used to define uncertainty are mean and variance (standard deviation) and many methods exist in graphical data analysis packages for displaying these quantities. This work extends these methods to include other descriptive statistics as well as distribution fitting techniques to not only provide visualizations of uncertainty, but also give the user a way to better understand the data distribution. Techniques for visualizing the correlation between multiple 1D categorical dataset are also presented (Figure 36), along with a variety of exemplar datasets.

## 4.8 Variable Interactions in Query-Driven Visualization

One fundamental element of scientific inquiry is discovering relationships, particularly the interactions between different variables in observed or simulated phenomena. Building upon our prior work in the field of Query-Driven Visualization, where visual data analysis processing is focused on subsets of large data deemed to be "scientifically interesting," this new work focuses on a novel knowledge discovery capability suitable for use with petascale class datasets. It enables visual presentation of the presence or absence of relationships (correlations) between variables in data subsets produced by Query-Driven methodologies. This technique holds great potential for enabling knowledge discovery from large and complex datasets currently emerging from SciDAC and INCITE projects. It is sufficiently generally to be applicable to any time of complex, time-varying,


Figure 37: Water ($H_2O$) and ethylene ($C_2H_4$) concentrations from the methane combustion dataset are shown in (a) and (b), respectively (Sample data courtesy J. Bell and M. Day, Center for Computational Sciences and Engineering, LBNL). The derived correlation field for these two compounds is shown in (c). The switch from strong positive correlation to strong negative correlation in the reaction region corresponds to the area in which $C_2H_4$ is both produced and consumed, and $H_2O$ is produced, in the process of combustion. The strong correlation (both positive and negative) in the center of the flame, as well as the atmospheric region, demonstrates the correlation field's ability to show fine-scale interactions.

multivariate data from structured, unstructured or adaptive grids.

Obstacles hindering scientific knowledge discovery from large and complex data may be broadly categorized into two separate but overlapping groups. The first category, concerned mainly with issues of throughput, includes the challenges inherent to efficiently managing and visualizing large-scale datasets. The second category includes the difficulties associated with attaining insight from datasets of high-complexity.

Query-driven visualization (QDV) is well suited for performing analysis and visualization on datasets that are both large and highly complex. Tools like FastBit leverage highly efficient (in terms of speed and compression) data management techniques to rapidly identify and visualize "regions of interest" within a dataset. Specified as Boolean range queries, these regions of interest tend to be significantly smaller subsets of the original dataset; thus, these regions require less time and effort to analyze, visualize, and interpret.

Well-characterized range queries are capable of identifying spatial regions where many domain-specific events occur: combustion flame fronts, vortices, chemical reaction fronts, etc. Beyond indicating these regions, however, queries reveal little about variable interactions or complex trends that lie in the domain of these characterizations. In such regions of interest, it is the behavioral trends between variables, or groups of variables, that are more important in providing insight than the traits or locations of individual variables alone. The challenge is to extend the strengths of QDV with methods that identify behavioral trends and provide insight into regions of interest through coherent and meaningful visualizations.

The novel contributions of this work are techniques that extend the capabilities of QDV by providing the basis for determining: how sets of variables in complex datasets interact throughout regions of interest, and the role other variables play in influencing these interactions.

We utilize the cumulative distribution functions (CDFs) of all variables in a query to reveal initial information about statistical regions of interest within the query's solution space. The CDF for each variable is computed by integrating over the query's solution space, then accumulating the variable's values as a histogram. Statistically, the solution set of a query is represented as an aggregate of histograms, one histogram for each variable expressed in the query.

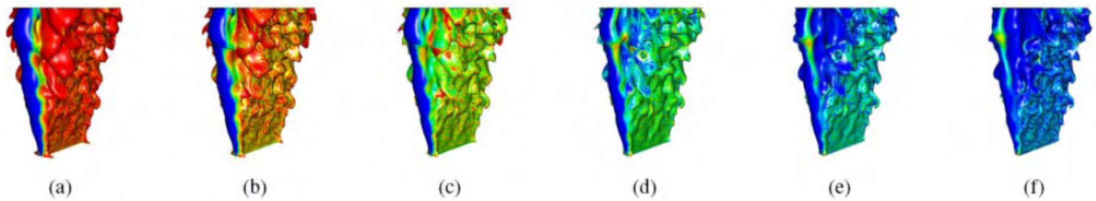



Figure 38: These images depict increase ((a) through (f)) isosurface values of temperature (isotherms) colored by values of the correlation field derived from water and ethylene. As temperature values increase, the predominant correlation between water and ethylene along the isotherms shifts away from strongly positive (red in (a)) to strongly negative (blue in (f)). This shift suggests that temperature is itself negatively correlated with the water-ethylene correlation. (Sample methane flame simulation data courtesy J. Bell and M. Day, Center for Computational Sciences and Engineering, LBNL)

We extend this analysis further by incorporating correlation fields, which provide insight into localized correlation between any two variables. By mapping a correlation field onto a third variable's isosurfaces (specifically, the statistically important isovalues suggested by the variable's CDF), statistically important interactions between any three variables in a dataset are readily visualized, allowing for trends between variables in a user's query to be identified.

In this method, CDFs and correlation fields are constrained to the query's solution space. By working exclusively in the query's solution space, this method takes full advantage of the performance benefits inherent to QDV strategies. Specifically, computational efforts are only focused on regions that have been rapidly identified (via a query engine) as "interesting" by the user's query. This method's integrated analysis extends current query solutions by revealing statistical trends of interactivity (i.e., dependency and independence) between any triad of variables in the solution space of the query.

# 5 Path Towards Petascale

As our highest-end computational platforms grow larger – moving through the petascale to the exascale and beyond – a forefront question that drives our work is "how do we best focus our efforts to support science at the petascale, and beyond?"

This question has several implications for our team. First is the notion of using these large machines for visual data analysis. Second is how to enable scientific knowledge discovery for the large datasets generated by simulations that machines, regardless of whether the analysis work is performed on the largest machine or on an auxiliary system, such as a large cluster, located near the large machine and intended for use as an analysis engine.

For the first issue, our team has worked to have our primary production-quality, parallel capable visual data analysis application, VisIt, included with the *Joule Metric*, which is a vehicle for quantitatively measuring and reporting computational efficiency gains on petascale-class platforms (Section 5.1). This work is "field-leading" because VisIt is the first-ever analysis or visualization application to play such a central role in the DOE HPC space.

For the second, we are taking a bifurcated strategy. We are providing direct and indirect support to numerous "petascale applications" (Section 5.2) as well as focusing our R&D effort on problems unique to scientific knowledge discovery of some of today's largest and most complex scientific datasets (Section 5.3).

## 5.1 Visualization on Petascale Platforms: VisIt and the Joule Metric

The Office of Management and Budget (OMB) uses its Performance Assessment Rating Tool (PART) to perform evaluations of different programs. PART has seven worksheets for seven types of agency functions, which includes R&D activities. In FY 2003, the DOE Office of Science worked directly with the OMB to come to a consensus on an appropriate set of performance metrics consistent with PART requirements. The scientific performance expectations of these requirements reach the scope of work conducted at the national laboratories. The *Joule* system emerged from this interaction. Joule enables the CFO and senior DOE management to track annual performance on a quarterly basis.

In FY09, OASCR's Joule goals include two primary performance metrics. The first concerns job size (e.g., level of concurrency) at NERSC. The second, and of interest in this discussion, concerns improving average performance increases in computational effectiveness (primarily through weak scaling studies).

In FY09, the VisIt visualization system is part of the Joule metric. VACET researchers from Oak Ridge National Laboratory are leading this effort. VisIt is the only visualization application that is included in the FY09 Joule metric.

For the Q2 benchmark, we have selected two common analysis and visualization techniques, isosurface extraction and volume rendering. These two algorithms will be run on the output of a recent simulation of the Denovo radiation transport code on the dose concentrations around a reactor core in a nuclear power generating plant. We demonstrate scaling in several ways. The first is weak scaling in extracting several different dose contours from the energy groups computed by the simulation code. Second, we demonstrate weak scaling in two aspects of the volume rendering: scaling with respect to the problem size, and scaling with respect to the number of samples computed along each ray.

In post-production analysis and visualization tools, the most important benchmarking metric is the time it takes to render a frame to a user display. Under normal uses cases, the user loads the simulation data from disk into memory and then repeatedly interacts with the data, successively modifying iso values and observing the results. We therefore focus on the scalability of the isosurface extraction algorithms and ignore the scalability of the onetime expense of reading simulation data from disk. As in isosurface extraction, the time to render frames once the data is loaded from disk is the most relevant metric to end-users. We therefore focus on the scalability of the volume rendering algorithms and ignore the scalability of the one-time expense of reading simulation data from disk.

The output produced from the VisIt analysis and simulation runs produced expected results that have been verified with the code developer and are deemed acceptable. We have therefore accepted the results for the Q2 benchmark runs.

> VACET researchers, by using VisIt as part of the FY09 Joule Metric, are helping DOE to meet its program objectives of showing increased computational efficiency on its leadership platforms.

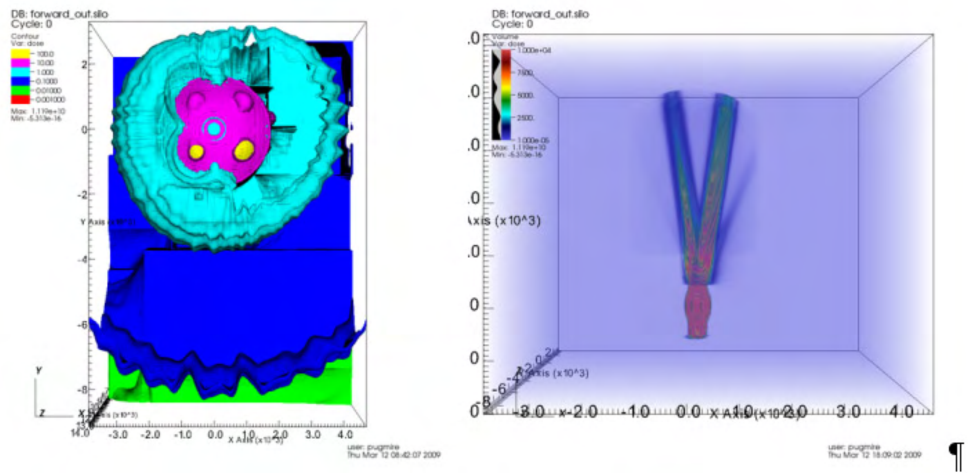



Figure 39: Extraction of radiation dose contours (left) and a volume rendering (right) from the nuclear power plant simulation results output from the Denovo radiation transport code.

## 5.2 Direct Support for Petascale Applications

A number of application codes have been selected as "Petascale Early Science" codes and awarded time at LCF/NCCS. This list is posted on the web[11]. From that list, we itemize those projects we are working with now (Table 4) or have either expressed interest in working with us or are already using VACET software (Table 5).

| Science Area | Project Name (Code) | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Astrophysics | A Three-Dimensional Model of SN987A (Chimera) | Bronson Messer and Tony Mezzacappa (ORNL) | • Develop and deploy petascale production turnkey visualization tool (VisIt) to interactively explore large time-varying datasets.<br>• Develop multidimensional visualization techniques for understanding complex radiation fields.<br>• Develop and deploy parallel streamlines algorithms for exploring magnetic fields.<br>• See Section 3.3 for more information about the work and Section 11 for Dr. Mezzacappa's testimonial. |

[11] http://www.nccs.gov/leadership-science/petascale-early-science/

| Science Area | Project Name (Code) | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Geoscience | Modeling Reactive Flows in Porous Media (PFLOTRAN) | Peter Lichtner (LANL) | • Developed a parallelized VisIt data reader for analyzing direct PFLOTRAN output.<br>• Modified VisIt TecPlot reader for reading PFLOTRAN legacy output formats.<br>• Several images for publication.<br>• Ongoing visualization support. |
| Materials Science | Investigations of the Hubbard Model with Disorder | Thomas Schulthess (ETH/ORNL) | • Supports 1/3 of an FTE of a VACET staff member to support visualization analysis needs.<br>• Ongoing visualization support. |
| Climate | Tests of Decadal Predictive Skill using the Community Climate System Model (CCSM) | Kate Evans (ORNL) | • Support for John Drake's climate team at ORNL through the use of Data Parallel R to understand rainfall variability and extreme weather events.<br>• Numerous images and movies from CCSM simulation runs for the climate team. |
| Biology | Cellulosic Ethanol: A Simulation Model of Lignocellulosic Biomass Deconstruction | Jeremy Smith (University of Tennessee) | • Production visualization support in the generation of publication-quality images. |

| Science Area | Project Name (Code) | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Nuclear Energy | Denovo, A Scalable HPC Transport Code for Multi-Scale Nuclear Energy Applications (Denovo) | Tom Evans (ORNL) | • Chosen as the primary source of data for the VisIt Joule metric scaling runs due to Denovo's scaling on Cray XT architectures and large domain counts.<br>• I/O support for complex data models (multiple energy groups, parallel ghost zone generation)<br>• Production visualization support for radiation dose exploration. |
| Combustion | Direct Numerical Simulation of Diesel Jet Flame Stabilization at High Pressure (S3D) | Jackie Chen (SNL-CA) | • Topological analysis R&D enables her to understand and see, for the first time ever, "non-visualizable" features: the characterization of extinction and reignition regions, and the role of transient autoiginition in the stabilization of lifted turbulent flames.<br>• Custom VisIt reader to directly read S3D output, allowing parallel analysis and visualization of S3D combustion runs. Deployed as part of the public VisIt distribution.<br>• Numerous images and animations.<br>• Chen hires staff formerly of part of VACET and students of VACET researchers.<br>• See Sections 3.8 for more information about the work and Section 11 for Dr. Chen's testimonial about our work. |

Table 4: Summary of Petascale Applications Receiving Direct VACET Support

| Science Area | Project Name (Code) | Collaborator | Summary of Accomplishments and Impact |
|---|---|---|---|
| Chemistry | Chemical Nanoscience at the Petascale (Unknown) | Robert Harrison (ORNL) | • PI has expressed interest in VisIt training so his team can generate their own analyses and visualizations. |
| Materials Science | Charge Patching Method for Electron Structures and Charge Transports of Organic and Organic/Inorganic Mixed Nanostructures (Unknown) | Lin-Wang Wang (LBNL) | • Wang's graduate student assistants have used VisIt (on their own) to generate images of 3D simulation data. |
| Nuclear Energy | Scalable Simulation of Neutron Transport in Fast Reactor Cores | Dinesh Kaushik (Argonne) | • The SHARP team uses VisIt and has an existing relationship with VACET personnel. |

Table 5: Summary of petascale applications already using VACET software or interested in VACET support.

## 5.3 Petascale Data and Scientific Knowledge Discovery

A natural consequence of simulations that run at increased spatial and temporal resolution, something made possible by machines with more cores and more memory, is an explosive growth in the size and complexity of simulation output. One of VACET's primary mission objectives is enabling scientific knowledge discovery in these increasingly challenging conditions. Our accomplishments in this area show good progress towards reaching this mission objective, enabling science at the petascale.

The following are representative of some VACET mission themes:

- **Provide reliable, scalable visual data exploration technology.** Our work with VisIt, not just for the Joule Metric, aims to ensure scalability to leverage bring to bear the capacity of parallel computing resources on ever-larger scientific datasets. Production-quality means that our science stakeholders have faith our applications are reliable and are willing to invest their resources into adopting it as well as "wagering" a portion of their projects' success on it.

> Out of the many analysis options where we may have chosen to concentrate our effort, VisIt stood out as particularly attractive because of the certainty it would scale. The *certainty* of this propsect was vital to our petascale planning as it allowed us to expend our contingency efforts in other areas. – *Paul Fischer, Argonne*

- **Accelerate scientific knowledge discovery.** In our query-driven visualization effort, we focus visualization and analysis processing, along with subsequent rendering and human cognition, on a subset of data deemed to be "scientifically interesting." This approach is akin to finding needles in haystacks, where the needle is specified with a compound (multivariate) boolean range query. Our work in this space leverages the award-winning FastBit index/query software from the SciDAC SDM Center, and has been applied to leading edge accelerator simulation code, hero-sized network traffic analysis problems, and combustion simulation data. For the accelerator project, we reduced the time required to find and analyze particles undergoing wakefield acceleration from hours to seconds/minutes. As data grows larger, these types of techniques become increasingly important – their runtime complexity is a function of the size of the selected subset rather than the size of the original data.

> Since it is impossible to explore and visualize very large datasets at once, innovative visualization technique have put to use the power of FastBit toq uickly select desired subsets of data. We have applied our joint technologies in explore accelerator design data, climate modeling data, network traffic data to identify malicious attacks, and others. . . . VACET has had a positive impact on our science effort, and is providing a valuable service to the scientific community. – *Arie Shoshani, LBNL*

- **Enable deeper scientific understanding than ever before possible.** Our work in topologically based feature detection and extraction (Sections 3.7, 3.8, and 8) reveal a common theme: a new regime of scientific insight is enabled through the use of advanced analysis techniques. These examples discuss new capabilities used to study difficult-to-quantify relationships in petascale-class data. In both cases, these scientists offer testimonials to the effect "for the first time, we can see" some aspect of their models that has to-date been elusive.

- **Directly address stakeholder large-data problems.** As FLOPS increase, so does the amount of output generated by simulations. This increasing amount of I/O spawns numerous challenges, a complete discussion of which is beyond the scope of this document. Suffice it to say that our team is addressing this challenge in a number of different ways. Our work with one of the SciDAC program's leading climate projects required our direct involvement in the design, implementation, testing, and optimization of a high performance I/O library: existing solutions, e.g. Parallel netCDF, did not meet this project's needs. Furthermore, long experience has shown that the best format/layout for writing data from simulations is not always the best for subsequent analysis operations. Finding a "happy medium" is also an ongoing challenge and is often best addressed on a project-by-project basis since there is no "one-size-fits-all" solution. This issue tends to receive attention during the analysis and visualization portion of the scientific investigation process primarily because most code teams have as a primary objective scalability. As a result, I/O concerns are secondary and often not addressed until late in the investigation.

# 6 Software Engineering

VACET must deliver production software to be successful. Production software allows the algorithms that VACET develops to affect the largest possible community, which is critical because the

demand from the SciDAC community for visualization and analysis capabilities far exceeds what a center of our size can deliver using only "one-off" solutions.

Production software requires rigorous software engineering: requirements gathering, design, development, testing, and maintenance. VACET embarks in all of these activities through its Software Engineering Group (SEG). This group has representation at each of VACET's five sites and it serves as the focal point for transitioning research software to production use. Its priorities are set by the Executive Committee based on feedback from its Customer Project Managers.

Building production software from scratch was infeasible giving VACET's funding, and also not necessary because of existing, good software. The SEG has gravitated towards VisIt as its dominant vehicle for delivering production visualization and analysis software to SciDAC stakeholders, because it is robust, well-documented and well-tested software, and, most importantly, it is positioned to handle petascale data. However, the SEG adapts its strategy to the situation, for example going away from the VisIt strategy by developing and incorporating SFSG (the Simple and Flexible Scene Graph) into the CDAT climate visualization tool, a preexisting and successful tool, enabling it to have three dimensional visualization capabilities.

The SEG has successfully transitioned many research efforts into production software, including VisIt's streamline algorithm, improvements to its volume renderer, incorporation of FastBit for query driven visualization, Poincaré plots, and others. Further, the SEG also adapts software to allow the SciDAC community to successfully use the tools provided by VACET. This often takes the form of writing file format readers for VisIt, adapting the user interface, adding small features, and fixing bugs. Each of these activities, although costly, is critical in user adoption of our software, which enables us to successfully deliver our research as well as provide visualization and analysis capabilities to the SciDAC community.

> VACET has made investments in VisIt that benefit the entire VisIt community, including ITAPS. They have invested in VisIt's software engineering infrastructure. . . . VACET's investment in VisIt provides momentum and leverage to a project that our center has a strong vested interest in. We appreciate their efforts . . . – *Lori Diachin, LLNL*

# 7 VACET Organization and Operations

## 7.1 Organization Overview

As described in the VACET proposal, our Center is organized into functional groups to achieve several distinct objectives: (1) facilitate the flow of information between the Center's leadership, personnel, and science stakeholders; (2) to provide the organizational structure needed to ensure oversight and coordinated operations of the Center's collection of activities; (3) to ensure we meet our work deliverables; (4) to gracefully accommodate future growth and respond to changing priorities. The Center's original functional groups are: the Center PIs, the Executive Committee, the External Advisory Board, Research and Development, Chief Software Engineer, Software Development and Support, and Stakeholder Projects. Due to a 45% reduction in budget, we have eliminated the formal External Advisory Board.

Figure 40 depicts the working interaction between these functional groups. The co-Directors (Bethel and Johnson) are responsible for the operation of the entire Center and are the point of contact between the Center and DOE. The Executive Committee (EC), which makes decisions about

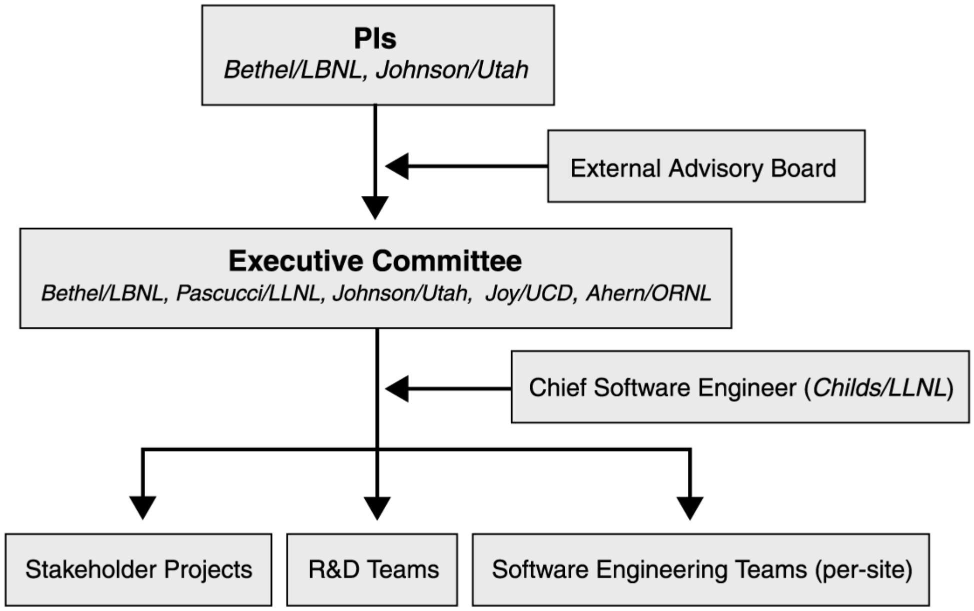


Figure 40: Functional interaction of groups within VACET.

research, development, stakeholder applications selection and prioritization, and software deployment, is comprised of the participating Laboratory or University PIs (Bethel, Johnson, Pascucci, Ahern, and Joy).

Given the high priority on delivering useful software to our scientific stakeholders, each lab site will have a primary software engineer whose duty it is to assure software developed and deployed from that site meets the Center software engineering criteria. The set of site engineers is known collectively as the Software Engineering Group (SEG). In addition to developing, testing, documenting and maintaining the Center software, the Software Engineering Group will integrate results from all research and development groups into the Center software, relying on feedback from the project leads in the Stakeholder Projects Group, the Executive Committee and the Chief Software Engineer (Childs).

The Chief Software Engineer (CSWE) serves many important functions within the center. One is to facilitate the coordinated design, implementation and integration of the Center's technologies into software solutions that meet stakeholder needs. He will provide guidance to the software development teams so that individual software tools and libraries will be readily usable throughout the Center's collection of stakeholder projects. He will coordinate with the EC to prioritize software development targets, and serve as a technical software adviser to the Center as a whole. He will interact with the R&D team to help foster early designs that fit well within the Center's technology implementations. He will direct the development, testing, deployment, and support of the Center software toolsets.

The Stakeholder Projects Group (SPG) is the primary interface to our science stakeholders. In this group, individuals from the Center will interact directly with science stakeholders to obtain and prioritize science needs, coordinate with the Center's EC and Chief Software Engineer to translate those needs into a work plan, to oversee and manage the work so that software is delivered to the science stakeholder.

From each participating institution is a "Site PI." That person is nominally part of the EC, and is responsible for managing issues at that organization. These issues range from responding to

administrative requests, like updated budgets, as well as for managing operations of VACET staff at that organization.

While our original proposal called for a team organization that includes an External Advisory Board (EAB), our project was funded at about 55% of the original request. Therefore, we have eliminated the formal EAB within the organization. Instead, we rely on direct feedback from stakeholders. After our first year of operation, this approach seems to be one that is viable, though not optimal.

### 7.2 Ongoing Operations: Internal Communications

The VACET management team (the Executive Committee and Chief Software Engineer) meet via teleconference at least once every two weeks. These discussions focus on issues ranging from high-level strategy, resource management, sharing information about progress in different program areas, project/resource prioritization, and so forth. These regular meetings are an indispensable part of ongoing management operations.

As a center, we meet formally twice per year in "All-Hands Meetings." The objectives for these meetings are: (1) for individuals to give presentations describing project progress; (2) for all hands to learn about other projects that are ongoing in the center; (3) to give more junior team members the opportunity to gain experience delivering technical presentations to a friendly yet demanding audience (there are always lots of questions and lots of discussion); (4) to create the opportunity for distributed teams to meet face-to-face to work together on joint projects; (5) to foster overall team spirit. We have conducted these meetings in the spring and fall so as to minimize conflict with known paper deadlines and to take advantage of events that many of us all likely to attend (e.g., IEEE Visualization).

### 7.3 Prioritization

Like many other projects, VACET is faced with having to make hard decisions about how to expend its resources. Naturally, we would prefer to have more resources to solve more problems. However, the overview of our strategy for making decisions about priorities is as follows: we rely heavily on input from our stakeholders to help us prioritize how we expend our precious R&D and software engineering resources. Two good examples are AMR visualization and flow visualization.

While we have "one" primary AMR visualization customer, the SciDAC Applied Partial Differential Equations Center (APDEC), that one stakeholder is responsible for nearly all AMR computational science R&D in DOE science programs. By working closely with APDEC, we are able, through a highly leveraged relationship, to effectively reach "all" of APDECs stakeholders, which include leading-edge efforts in fusion, combustion, and computational astrophysics.

Early in the VACET project, we solicited needs from our science stakeholders. Nearly all of them mentioned the need for better flow visualization algorithms and infrastructure. We have responded with a focused R&D effort that (1) has produced an excellent, accurate, stable, and robust solution to this problem that is applicable to a family of flow visualization algorithms (streamlines, particle advection, stream surfaces) and derivative algorithms (Poincaré plots); and (2) is deployed in and delivered to the world in VisIt, so as to be able to run in parallel on basically all modern HPC platforms.

Essentially all R&D activities in which we engage can be traced back to stakeholder needs, which are accumulated in a private area on the VACET wiki. Furthermore, this approach helps focus our efforts on those that are likely to have a large payout in terms of reaching a large user base and that will actually be of significant benefit to them.

# 8 Two-Page Science Highlights

# Accelerating Accelerator Science

O. Rübel[1], Prabhat[1], D. Ushizima[1], G. Weber[1], J. Jacobsen[1], C. G. R. Geddes[1], K. Wu[1], H. Childs[2], J. Meredith[3], E. Cormier-Michel[1], S. Ahern[3], P. Messmer[4], E. W. Bethel[1]
[1]Lawrence Berkeley National Laboratory; [2]Lawrence Livermore National Laboratory; [3]Oak Ridge National Laboratory; [4]Tech-X Corporation

## Summary

Scientists from the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) have developed new capabilities that "accelerate accelerator science" by enabling rapid scientific knowledge discovery. These new techniques leverage DOE's parallel computing platforms to: (1) reduce from hours to seconds the time required to select particles undergoing acceleration in simulation results and to track them over time; (2) automatically locate particles undergoing acceleration and to analyze them across time and space through the use of advanced machine learning techniques. These new capabilities represent a major step forward for DOE's Accelerator research programs.

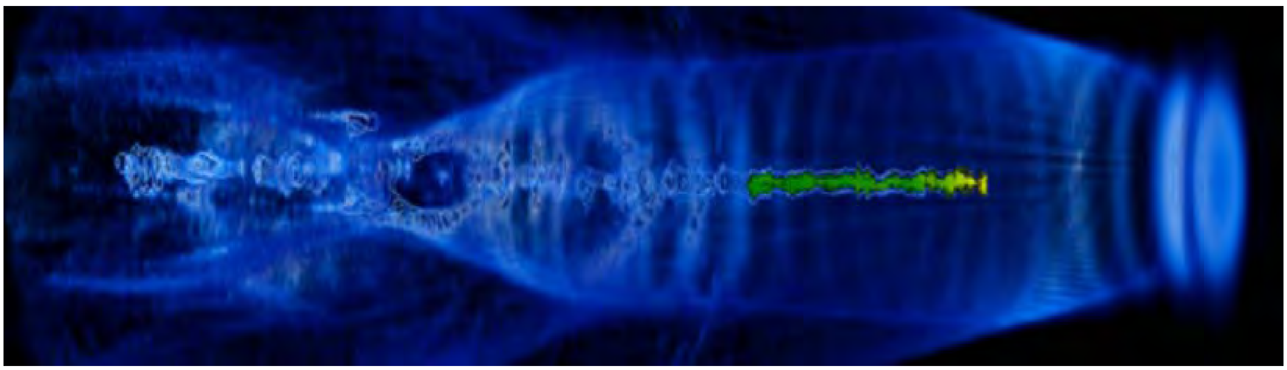


**Figure 1.** Petascale-class machines help accelerator scientists to model and plan experiments with better fidelity and physical accuracy then ever before possible. Output from one such simulation is shown here with the wake (blue) and beam electrons (yellow), which show a high degree of acceleration and spatial coherency. Image courtesy C. Geddes, LBNL.

In laser wakefield accelerators, particles are accelerated to relativistic speeds upon being "trapped" by the electric fields of plasma density waves generated by the radiation pressure of an intense laser pulse fired into a plasma. These devices are of scientific interest because they are able to achieve very high particle energies within a relatively short amount of distance when compared to traditional electromagnetic accelerators. Advanced accelerator designs, include laser wakefield accelerators, is one component of the portfolio in the SciDAC Science Application *Community Petascale Project for Accelerator Science and Simulation (ComPASS)*.

Researchers use computer simulation to model experiments and test designs for future experimental devices. Petascale class machines enable simulation runs at unprecedented levels of spatial resolution and physical accuracy. One challenge facing this scientific research is the sheer volume and complexity of data being produced by simulations.

The SciDAC *Visualization and Analytics Center for Enabling Technology (VACET)* has developed new technologies, now in the hands of scientific researchers, that accelerate scientific knowledge discovery by: (1) reduce from hours to seconds or minutes the time required for finding accelerated particles in simulation results and tracking them over time; (2) replacing a manual particle search process with one that is automated and based upon state-of-the-art machine learning. Together, these two capabilities help to increase

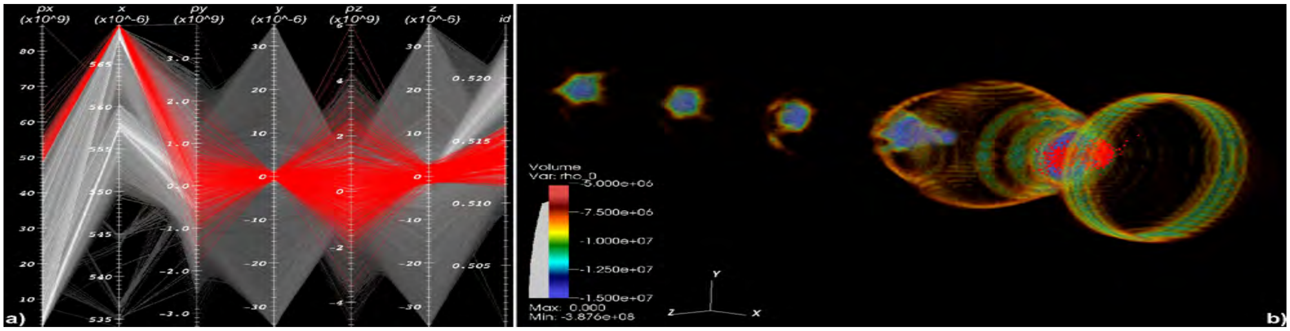



**Figure 2.** The interface on the left, which shows a multivariate view of particles' distribution across all simulation variables, is the vehicle for selecting particles having high acceleration and spatial coherency. All particles are shown in gray, and selected particles shown in red. This selection interface is linked with other forms of visual data exploration, so that a physical view of selected particles (right, red particles) helps the scientist to quickly gain insight into the relationship between statistical-space and physical-space features. This new capability, developed as part of VACET's research and development portfolio, is now part of widely distributed, production-quality, parallel capable visual data exploration software infrastructure (VisIt). Image courtesy O. Rübel et al., LBNL.

the rate and quality of scientific knowledge discovery.

Working with researchers from the SciDAC *Scientific Data Management Center*, the VACET team developed a new capability for the rapid visual exploration of very large, multivariate, time-varying datasets. Using this new technology, a scientist interactively selects subsets of data via a user interface that presents a statistical distribution (Figure 2, left) of particle information across all simulation dimensions, e.g., position, velocity, electrical charge, etc. Concurrent with exploration in statistical space, the subset is simultaneously displayed in physical space (Figure 2, right). Once a scientist finds interesting particles in one time step, the new technology will locate all those particles across all timesteps. Further investigations of this complete set of particles helps to increase scientific understanding of the processes that lead to wakefield acceleration.

These new capabilities are implemented in production-quality, parallel-capable visual data exploration software that runs on virtually all modern platforms, ranging from desktop-class machines to DOE's petascale computer systems.

Both these steps represent a major new capability for accelerator scientists: before this technology, the search and tracking process consumed many hours of computer time; now, these processes take only a few seconds, even with today's largest accelerator simulation datasets, when run on DOE's parallel computing platforms.

To help further accelerate scientific knowledge discovery, VACET researchers developed new technology for automatically finding and tracking particles of interested, those undergoing wakefield acceleration. The idea is to use machine learning technology to find particles undergoing rapid acceleration and a high degree of spatial coherence. Once located, these paths of these particle "bunches" can be quickly displayed and analyzed. This new capability helps to accelerate scientific knowledge discovery by replacing a manual search process with one that is automated.

**Recent Publications**

O. Rübel, Prabhat, K. Wu, H. Childs, J. Meredith, C.G.R. Geddes, E. Cormier-Michel, S. Ahern, G.H. Weber, P. Messmer, H. Hagen, B. Hamann and E.W. Bethel, "High Performance Multivariate Visual Data Exploration for Extremely Large Data." SC08, Austin TX, November, 2008. LBNL-716E.

Daniela Ushizima, Oliver Rübel, Prabhat, Gunther Weber, E. Wes Bethel, Cecilia Aragon, Cameron Geddes, Estelle Cormier-Michel, Bernd Hamann, Peter Messmer, Hans Hagen. "Automated Analysis for Detecting Beams in Laser Wakefield Simulations." 2008 Seventh International Conference on Machine Learning and Applications, Proceedings of IEEE ICMLA'08, 2008. LBNL-960E

**For further information on this subject contact:**

Name: E. Wes Bethel.
Organization: Lawrence Berkeley National Laboratory.
Email: ewbethel@lbl.gov
Phone: (510) 486-7353

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# Novel Visualization Algorithms and Production Visualization Tools for Adaptive Mesh Refinement Data

G. H. Weber[1], H. Childs[2], E. Deines[3], C. Garth[3], J. Meredith[4], D. Pugmire[4], B. Whitlock[2], E. Wes Bethel[1]

[1]Lawrence Berkeley National Laboratory; [2]Lawrence Livermore National Laboratory; [3]University of California, Davis; [4]Oak Ridge National Laboratory

## Summary

Adaptive Mesh Refinement (AMR) is a highly effective simulation method for spanning a large range of spatiotemporal scales like those encountered in astrophysical simulations that must accommodate ranges from interstellar to sub-planetary. Historically, AMR data have not been supported well, if at all, in visual data exploration applications. As a result, AMR code teams had no choice but to maintain in-house applications, such as ChomboVis, for AMR visualization. The Department of Energy's (DOE's) Science Discovery through Advanced Computing (SciDAC) Visualization and Analytics Center for Enabling Technologies (VACET) has extended VisIt, an open source visualization tool that accommodates AMR as a first-class data type, in an effort that consists of research in novel AMR visualization algorithms, such as AMR streamlines, as well as visualization infrastructure engineering. As a result, VisIt is now the premier, production-quality, parallel-capable AMR visual data analysis tool. The Applied SciDAC Applied Partial Differential Equations Center for Enabling Technologies (APDEC) and its collaborators have adopted VisIt for all their AMR visualization needs, supplanting ChomboVis, their in-house application, resulting in significant cost and labor savings.

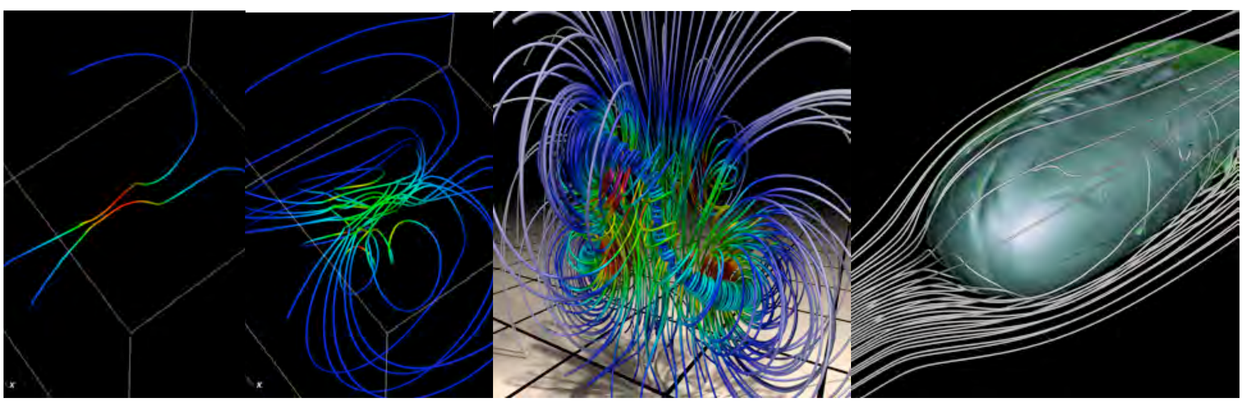


**Figure 1: (**Left) Calculating streamlines without considering the AMR hierarchy does not detect finer AMR levels that represent an area within the domain at higher accuracy. Instead, streamlines stay completely in the coarsest level. Considering the AMR hierarchy during streamline calculations makes it possible to detect the presence of finer resolution levels and always use the most accurate data representation available. (Center) Production-quality visualization of an AMR vortex core merger simulation for a brochure. (Data courtesy of D.Martin, APDEC.) (Right) Using streamlines for visualizing an AMR solar wind simulation. (Data courtesy of S. Borovikov, University of Alabama, Huntsville.)

Adaptive Mesh Refinement (AMR) techniques combine the compact, implicit structure of regular, rectilinear grids with the adaptivity to changes in scale of unstructured grids. VisIt is an open source visualization tool that accommodates AMR as first class data type. VisIt offers a rich set of production-quality functions for parallel visualization and analysis of complex data sets on

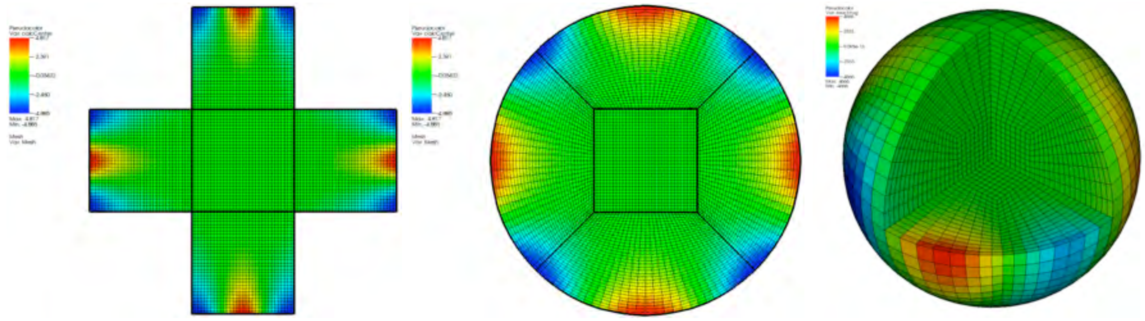


**Figure 2:** VisIt's mapped grid support for AMR data lets users switch between parameter space (left) and physical space (center). It supports both 2D (left, center) and 3D (right) mapped AMR grids, which is important for fusion and climate science applications. (Data courtesy of D. Martin, APDEC.)

parallel platforms, making it an ideal candidate to replace specialized AMR visualization tools.

AMR visualization is challenging since coarser information in regions covered by finer patches is superseded and replaced with information from these finer patches. During visualization, it becomes necessary to manage selection of which resolutions are being used for any given visualization operation. Furthermore, it is difficult to avoid discontinuities at level boundaries, which, if not properly handled, lead to visible artifacts in visualizations.

Among our AMR visualization research efforts is a novel streamline calculation algorithm. Working closely with our APDEC stakeholders, we are developing AMR-aware streamlines in VisIt (Figure 1). As outlined above, the major algorithmic challenge lies in detecting the presence of finer AMR hierarchy levels, ensuring that streamlines are computed with the highest available resolution data, and using correct interpolation at the transition between coarse and fine levels. The involvement of APDEC researchers in this project ensures that streamline calculations meet stringent accuracy requirements.

Our AMR visualization work is also focused on specific application domains – fusion and computational astrophysics who are APDEC stakeholders – to ensure that the new algorithms we have and continue to develop meet their needs. VisIt serves as our primary research and development infrastructure for parallel AMR visualization. As such, new research quickly transitions from prototype to deployment in a production-quality end-user tool.

We are also leveraging VisIt's capability to deform a computational grid or grid hierarchy by a vector variable. This ability makes it possible to view simulation data both in computational parameter space as well as in "mapped" physical space. Working with APDEC researchers, we extended VisIt's Chombo file reader to test for the existence of mapping information. Further, we added macros that simplify accessing this functionality and allow users to switch seamlessly between computational space and physical space views (Figure 2). Switching between these views is essential for debugging purposes, since it supports viewing both the logical grid structure of simulations, such as the connectivity between the five grids on the left side of Figure 2, as well as seeing results in a physically meaningful context.

Close collaboration with APDEC allows us to develop this visualization capability in conjunction with the corresponding simulation codes, ensuring that appropriate visualization capabilities are deployed shortly after simulation code release. The ability to handle mapped grids is important to application scientists in climate (e.g., the cubed sphere on the right hand side of Figure 2) and fusion applications, who were recently able to view their AMR simulation results in mapped space for the very first time.

**For further information on this subject contact:**
Name: Gunther H. Weber.
Organization: Lawrence Berkeley National Laboratory.
Email: GHWeber@lbl.gov
Phone: (510) 486-5871

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# Understanding Pre-Mixed Turbulent Combustion

Peer-Timo Bremer[1], Gunther Weber[2], Valerio Pascucci[3], Marc Day[2], John B. Bell[2]
[1]Lawrence Livermore National Laboratory; [2]Lawrence Berkeley National Laboratory; [3]University of Utah

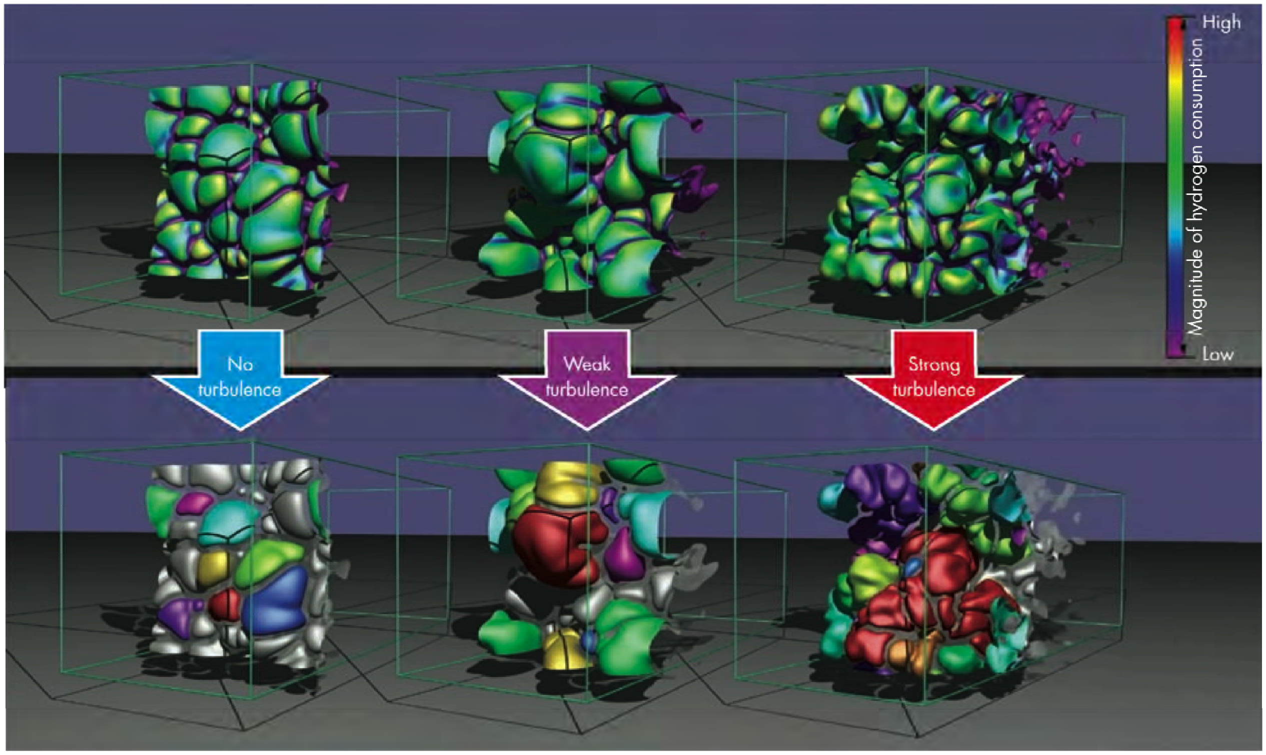
**Summary**

There have been significant recent efforts to better understand the nature of turbulent combustion in order to guide the design of new fuel-efficient, low-emission engines and turbines. However, adding turbulence to the already complicated combustion process results in seemingly chaotic intricate reactions that are difficult to analyze. Scientists from the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) have developed a suite of general purpose data analysis tools based on Morse theory that can reliably process large, time-dependent scientific data sets regardless of their underlying complexity. Using Morse theory, we compute hierarchical topological graphs that encode complete families of structural segmentations. Despite being orders of magnitude smaller than the original data, these graphs enable data exploration at multiple spatiotemporal scales and within a wide range of parameters. In particular, the hierarchy supports extensive parameter studies without the need for potentially expensive re-processing of the original data. The resulting statistics enable scientists to select appropriate parameters and provide insight into the impact of the choice of parameters on the results. These new methods enable a new form of scientific understanding: the ability to quantitatively correlate the turbulence of the burning process with the distribution of burning regions properly segmented and selected. Our analysis shows a new, counter-intuitive result: stronger turbulence leads to larger cell structures, which burn more intensely than expected.

Understanding combustion processes over a broad range of operational regimes is of great interest for a variety of applications like engine or power plant design. To this end, there has been considerable recent interest in the development of premixed burners capable of stably burning ultra-lean hydrogen-air fuel mixtures. Such burners could, for example, be used as one component of a clean-coal power plant utilizing hydrogen extracted from coal gasification. Lean premixed systems are subject to a variety of hydrodynamic and combustion instabilities that render practical flame stabilization, and traditional approaches to flame analysis, extremely difficult. The flames burn in a cellular mode that is highly non-uniform, time-dependent, and difficult to characterize.

To study the combustion process at different levels of turbulence, scientists at the Lawrence Berkeley National Laboratory performed numerical simulations of lean premixed hydrogen flames in three different turbulence configurations: no turbulence, weak turbulence, and strong turbulence. Our analysis characterizes the burning behavior in these systems by the number and size of burning cells. We represent flames as isotherms of constant temperature. On each flame surface, thresholds on the local fuel consumption divide the surfaces into burning cells separated by non-burning regions.

To gain new insights into the combustion process, scientists are interested in two types of analysis: time-aggregate statistics and detailed analysis of cell evolution. Comprehensive statistics on the number and area of burning cells aggregated over time provide important information about quantitative and qualitative differences between flames under different levels of turbulence. A tracking graph representing the evolution of cells over time describes the local cell dynamics in detail and enables an in depth study of all temporal events, such as cell births, deaths, merges, and splits. Coupling the tracking graph to an interactive visualization of the segmented flame surfaces illustrates each event and is an important tool to verify and validate the parameter selection.



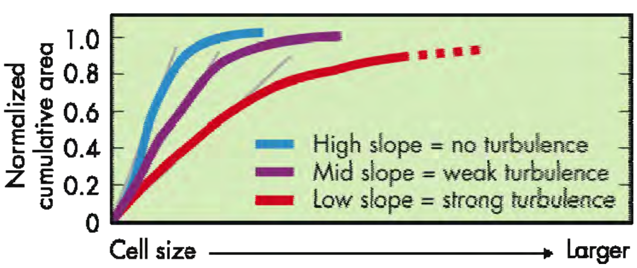




**Figure 1.** *The top 3D diagrams show flame surfaces of a lean premixed hydrogen flame at different levels of turbulence colored by the local fuel consumption. In the bottom 3D diagrams a small set of burning cells are randomly colored to show the irregularity of the more turbulent cells. In the graph, the corresponding cumulative density function of cell area distributions show that more turbulence creates larger cells with a wider distribution of normalized surface areas indicating a more intense burning process.*

However, no unique correct threshold exists, so a primary goal is to determine and verify the choice of parameters. Using the parallel computing resources at NERC we extract the isotherms for each time-step of the simulation and compute merge trees of the fuel-consumption function defined on the isotherms. Each tree encodes the cell configurations for all possible thresholds, and we augment the trees by storing cell areas as a function of the threshold along their corresponding branches. We use the augmented trees to perform extensive parameter studies, which determined a viable threshold and demonstrated the consistency of the results using varying parameters. Once we determine a threshold, we track the corresponding burning cells through time and create a tracking graph that encodes their temporal evolution. Our methods allow, for the first time, a quantitative analysis of the cellular burning structures and yield important scientific insights: Although it seems counterintuitive, higher turbulence levels lead to larger cell structures, which also burn more intensely than predicted by simple theories of flame propagation. These results suggest that premixed hydrogen flames could be stabilized at much leaner conditions than previously believed.

**Recent Publications**

M. Day, J. Bell, P.-T. Bremer, V. Pascucci, V. Beckner, M. Lijewski, "Turbulence effects on cellular burning structures in lean premixed hydrogen flames", *Combustion and Flame*, in press.

G. Weber, P.-T. Bremer, V. Pascucci, M. Day, and J. Bell, "Feature Tracking Using Reeb Graphs", in *Proc. 3rd TopoInVis Workshop*, to appear.

**For further information on this subject contact:**
Name: Peer-Timo Bremer.
Organization: Lawrence Livermore National Laboratory.
Email: bremer5@llnl.gov
Phone: (925) 454-0465

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and has been performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

# Advancing Astrophysics Science

D. Pugmire[1], R. Toedte[1], B. Messer[1], S. Ahern[1], E. Endeve[1,2], R. Budiardja[1,2]
[1]Oak Ridge National Laboratory; [2]University of Tennessee, Knoxville

**Summary**

Petascale simulations of the deaths of massive stars are important to the understanding of the origin of the diversity of elements within the universe. The data sets from these petascale simulations are becoming larger due to increases in spatial and temporal resolution as well as large increases in data model complexity. The SciDAC Visualization and Analytics Center for Enabling Technology (VACET) has developed new techniques that enable rapid scientific knowledge discovery from these data sets. These techniques leverage the parallel data analysis platforms at DOE's supercomputing centers to provide "turn key" visualization and analysis for DOE's astrophysics research programs.

Astrophysics researchers from Oak Ridge National Laboratory, Florida Atlantic University, and North Carolina State University are working, through DOE's INCITE program, to develop petascale simulation codes with the capability of modeling the complex physics of core-collapse supernovae. Such phenomena are arguably the most important links in our chain of origin from the Big Bang to the present day. They are the dominant source of elements in the periodic table between oxygen and iron and are responsible for producing half of the elements heavier than iron. Ascertaining the explosion mechanism for core-collapse supernovae is one of the most important unsolved problems in astrophysics.

INCITE researchers are developing a fully coupled, extensible radiation hydrodynamic code, CHIMERA, that implements scalable multi-dimensional hydrodynamics and magnetohydrodynamics solvers and neutrino radiation transport solvers, along with nuclear kinetics solvers and gravitational solvers for both Newtonian and general relativistic gravity.

A core-collapse supernova is initiated by the collapse of the core of a massive star. The core rebounds at high density, launching a shock wave into the star that will ultimately disrupt it. The shock stalls in the core, however, losing energy as it plows through the still infalling stellar matter. Exactly how the shock is revived is still unknown. This is the central question in core-collapse supernova theory today. Core-collapse supernovae are in part radiatively driven. After core bounce, $10^{53}$ ergs of energy in the form of neutrinos of all

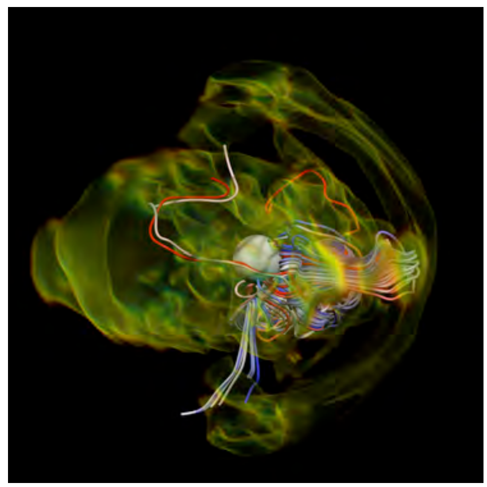


**Figure 1.** The collapse of massive star's core results in the formation of an outgoing spherical shock wave that eventually disrupts the entire star, giving rise to a supernova. Along the way the shock temporarily stalls and experiences the "stationary accretion shock instability" (SASI), which causes large deviations from spherical symmetry. This appears to be important to the supernova explosion mechanism, and may be responsible for spinning up the collapsed core – a nascent neutron star – into a pulsar. This image shows an exploratory view of a simulation run to ascertain the extent to which the SASI may generate magnetic fields: a volume rendering shows the fluid speed, and a sampling of fluid streamlines is colored by magnetic field strength. The simulation was run on Jaguar at NCCS with *GenASiS*, a multi-physics code under development for the simulation of astrophysical systems involving nuclear matter. Image credit: Visualization: Dave Pugmire (Oak Ridge National Laboratory). Simulation: Eirik Endeve, Christian Cardall, and Reuben Budiardja (Oak Ridge National Laboratory and University of Tennessee, Knoxville)

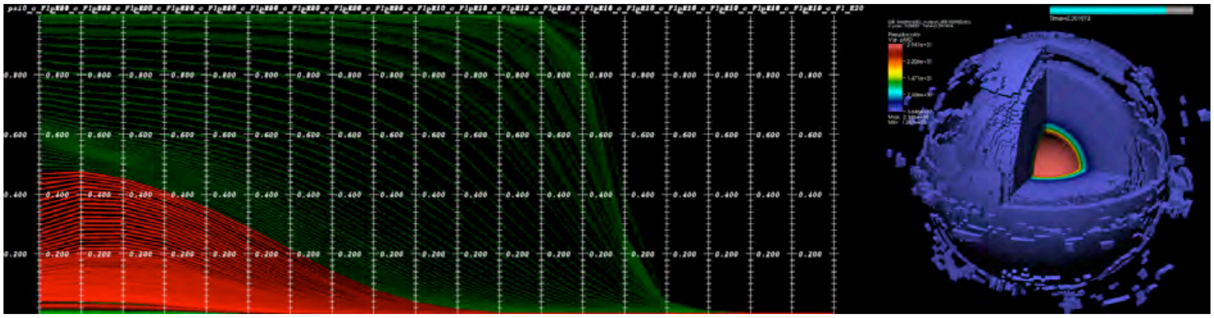


**Figure 2.** Multidimensional parallel coordinates exploration of the 20-energy group neutrino radiation field in a large supernova simulation. Increasing neutrino occupation highlighted in red. The multidimensional analysis on the left is directly linked to a traditional spatial visualization on the right. Visualization: Sean Ahern (Oak Ridge National Laboratory). Simulation: Steve Bruenn (FAU), Bronson Messer, Tony Mezzacappa, Raph Hix (Oak Ridge National Laboratory)

three "flavors" (electron, mu, and tau) is released from the "proto-neutron star (PNS)" at the center of the explosion. The supernova explosion energy is $10^{51}$ ergs, one hundred times smaller. Energy in the form of neutrinos emerging from the PNS will be deposited behind the shock and will help revive it. This "neutrino reheating'" is central to the core-collapse supernova mechanism. However, while a prodigious amount of neutrino energy emerges from the PNS, the neutrino heating is very sensitive to the distribution of neutrinos in energy (i.e., their spectra) and direction of propagation (i.e., their angular distributions) at any given spatial point behind the shock. Given that different neutrino flavors have different spectra, mixing between flavors may be crucial. Thus, multi-angle, multi-frequency neutrino transport is ultimately required to accurately compute the neutrino distributions in the heating region.

All of this complexity renders the core-collapse supernova problem a truly multidimensional, petascale problem, and it demands analysis and visualization tools that can be brought to bear in taming such complexity. The CHIMERA team has turned to the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) for assistance and has increasingly relied upon VACET's VisIt parallel analysis and visualization tool for understanding their petascale datasets.

Specifically, the capabilities for large-scale data parallelism in the generation of images and analyses for exploring pressure, density, and entropy fields have been deployed into the hands of scientists.

In the past two years, VACET staff members have added multi-dimensional analysis techniques that allow for scientists to directly explore the complex data models inherent in their files. These techniques have, for the first time, allowed the researchers to directly see the energy distribution and propagation direction of neutrino radiation in their simulations. In addition, these multi-dimensional visualizations are linked to traditional spatial visualizations, allowing for focused analysis of features of interest.

These new capabilities are implemented in production-quality, parallel-capable visual data exploration software that runs on virtually all modern platforms, ranging from DOE's largest petascale computer systems down to laptop-class machines.

**For further information on this subject contact:**

Name: Sean Ahern
Organization: Oak Ridge National Laboratory.
Email: ahern@ornl.gov
Phone: (865) 241-3748

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

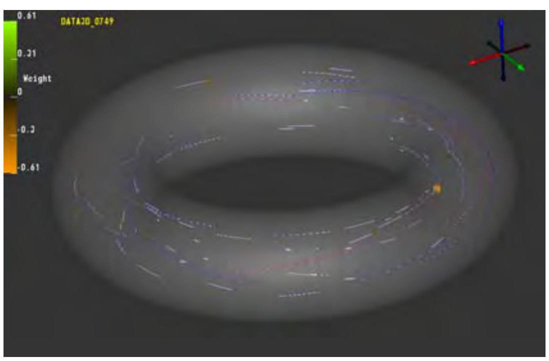
# Rapid Analysis of Plasma Instabilities in Fusion Science

Allen R. Sanderson[1], Prabhat[2], O. Rübel[2], G. Weber[2], H. Childs[2], J. Meredith[3], S. Ahern[3], David Pugmire[3]

[1]University of Utah; [2]Lawrence Berkeley National Laboratory; [3]Lawrence Livermore National Laboratory; [3]Oak Ridge National Laboratory

## Summary

Developing a reliable energy system that is economically and environmentally sustainable is the long-term goal of Fusion Energy Science (FES) research and is a worldwide effort. As fusion experiments have increased in size and complexity (too expensive to duplicate), there has been concurrent growth in the importance of simulation and visualization. Scientists from the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) working closely with fusion scientists have developed new capabilities for visually exploring and, ultimately, understanding fusion simulation data sets. These new techniques leverage DOE's parallel computing platforms enabling the: (1) rapid exploration of datasets containing millions to billions of values, (2) rapid extraction of features of interest within a large simulation, and (3) interactive visualization of the analysis results. These techniques have helped fusion scientists reduce the time spent in analysis and further the utility of fusion reactor simulations.



Output from microturbulence simulation of a fusion plasma showing 22 of the 400 million plasma particles simulated. Petascale-class machines help fusion scientists model their simulations with higher fidelity and physical accuracy then previously possible. Particles are displayed based on the number of times that they were magnetically trapped (red line) and de-trapped (blue line) in relation the confining magnetic field. Data courtesy S. Ethier, PPPL.

A critical characteristic of a typical fusion reactor is the growth of instabilities in the plasma due to the large gradients of density and temperature, the field geometry, and the inherent self-consistent interactions between charged particles and electromagnetic waves. Identifying and understanding these plasma instabilities is critical in the design of fusion reactors such as the International Thermonuclear Experimental Reactor (ITER), a Tokomak reactor scheduled for completion in 2018. Plasma instabilities occur on very different spatial and temporal scales and can represent highly unique phenomena.

For instance, fusion scientists are actively exploring the micro-scale instabilities created by microturbulence within the plasma. The waves generated by this turbulence can trap particles within the electric field, resulting in a radial diffusion of plasma particles across the confining magnetic field. On the macro scale fusion scientists are interested in "magnetic island" formation due to anisotropic plasma equilibration relative to the magnetic field. As with microtrubulence, magnetic island formation allows plasma particles to escape from the confining magnetic field. Understanding these two phenomena is necessary in the ultimate understanding and control of the rate of energy loss and plasma deposition on the wall of the fusion reactor.

The SciDAC VACET team has developed new technologies that are currently in use in the fusion community. The software deployed enables; (1) the rapid finding and tracking of a plasma particle within a given simulation (previously this task required hours to complete

and now requires only seconds); (2) the replacement of a manual particle search process with one that is automated and based upon state-of-the-art machine learning, and (3) the rapid identification of “magnetic islands” within the magnetic field.

Working with researchers from the SciDAC *Scientific Data Management Center*, the VACET team developed a method for the rapid visual exploration of very large, multivariate, time-varying datasets. Initially developed for analyzing particles in laser wakefield simulations, this method was expanded to meet the needs of fusion scientists. This expanded capability includes the ability to perform explorations and track particles on a cumulative basis as well as utilizing derived data.

The VACET team, in collaboration with the SciDAC Center for Extended Magnetohydrodynamic (MHD) Modeling, has developed analysis tools for automatically identifying and extracting magnetic islands. This analysis tool, which has its foundation in time–frequency techniques, allows fusion scientists to either work directly within their simulation codes using portable libraries or work offline after a simulation has completed.

The new capabilities provided by the VACET team to the fusion research community are implemented in production-quality, parallel-capable, visual data exploration software that runs on virtually all modern platforms, ranging from desktop-class machines to DOE’s petascale computer systems.

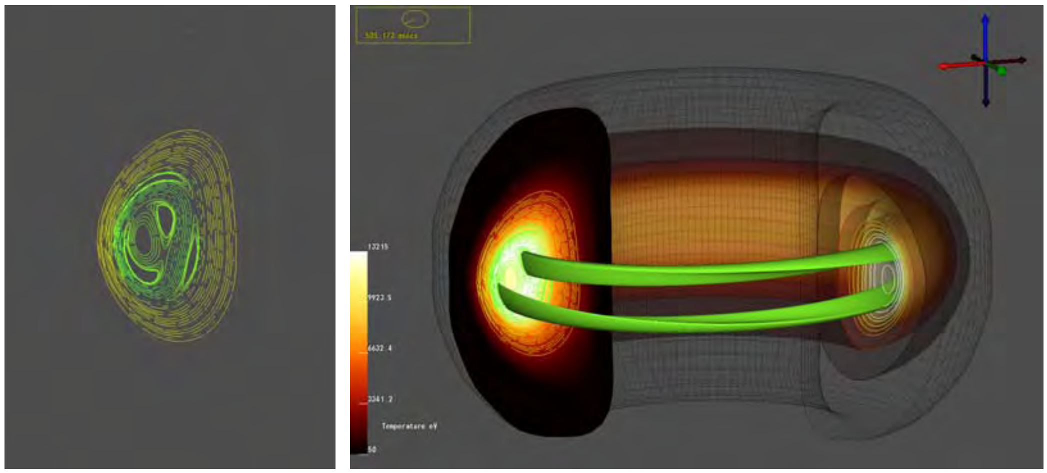


**Figure 2:** The topology of the magnetic field is visualized using an analysis tool that produces a Poincaré map, left. Of particular interest in the analysis of the magnetic field topology is the break up of the magnetic flux surfaces into a series of “magnetic island” chains. The image on the right shows an island chain that dominates the inner core of the simulation. As the field becomes more stochastic, the plasma will cool rapidly. This cooling is highlighted by a series of transparent iso-temperature surfaces. Though the temperature profile remains as a series on nested contours they have deformed based on the topology of the magnetic field. This visualization enables scientists to quickly gain insight into the relationship between statistical-space and physical-space features. This tool, developed as part of VACET’s research and development portfolio, is being deployed within a widely distributed, production-quality, parallel capable visual data exploration software infrastructure (VisIt).

**Recent Publications**

Allen R. Sanderson, Xavier Tricoche, Christoph Garth, Scott Kruger, Carl Sovinec, Eric Held, and Joshua Breslau, “Detection of Magnetic Nulls in Toroidal Geometry”, 48th Annual Meeting of the Division of Plasma Physics, 2006.

Allen R. Sanderson, Xavier Tricoche, Christoph Garth, Scott Kruger, Carl Sovinec, Eric Held, and Joshua Breslau, “Visualizing Patterns in the Poincare Plot of a Magnetic Field”, IEEE Visualization Conference Compendium 2006.

**For further information on this subject contact:**

Name: Allen R. Sanderson
Organization: SCI Institute, University of Utah.
Email: allen@sci.utah.edu
Phone: (801) 585-0769

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR’s Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# Massive, Distributed and Highly Accurate Computation of Particle Trajectories

D. Pugmire[1], S. Ahern[1], H. Childs[2], G. Weber[3], C. Garth[4], E. Deines[4], K. I. Joy[4]

[1]Oak Ridge National Laboratory; [2]Lawrence Livermore National Laboratory; [3]Lawrence Berkeley National Laboratory; [4]Institute for Data Analysis & Visualization, University of California, Davis

## Summary

Scientists from the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) have developed new capabilities for the analysis and visualization of very large vector field datasets with the aim of enabling rapid scientific knowledge discovery through state-of-the-art visualization methods. These new techniques leverage parallel computing platforms to significantly reduce the computational time required for computing particle trajectories in vector fields output from petascale simulations. Recent advances enable correct and reliable computation of particle trajectories in very large Adaptive Mesh Refinement (AMR) multi-grid datasets.

**Figure 1.** A path surface illustrates the flow of air around a car. The surface represents a dense sheet of massless particles, and the surface color coding provides insight into the lifetime of individual particles. The most notable features identified in this visualization are the vortices behind the rear-view mirror, where the particles are drawn into a vortex (close-up on the right). The path surface is constructed from several thousand integral curves.

Modern methods for the analysis and visualization of vector fields, which play a key role in many application domains like flow simulation, astrophysics and fusion, are built on empirically studying the behavior of massless particles advected with the vector field. These so-called *Lagrangian* methods offer unparalleled insight into vector field structures especially for time-dependent vector fields.

One simple, but quite effective, visualization approach for vector fields is the straightforward depiction of particle trajectories, as described by vector field integral curves and computed with the aid of numerical integration. Recently, VACET scientists have developed a novel approach of grouping trajectories of particles emanating from a common curve into so-called stream and path surfaces. The resulting visualizations – stream surfaces – are a significant improvement over visualizing individual curves due to the fact they offer the possibility for greatly improved visual understanding of complex 3D flow features.

Moreover, the recently introduced notion of *Finite-Time Lyapunov Exponents* (FTLE) has garnered much attention in both visualization and application science communities for its ability to visualize and analyze time-varying vector fields in terms of the convergence and divergence of neighboring particles. Having obtained such information, it is possible to accurately depict the fully dynamic nature of time varying vector fields. In contrast, earlier techniques are unable to easily capture time-varying structures. In addition, FTLE enables the representation of the structural dynamics in terms of scalar fields, opening up vector field structural analysis to a host of ex-

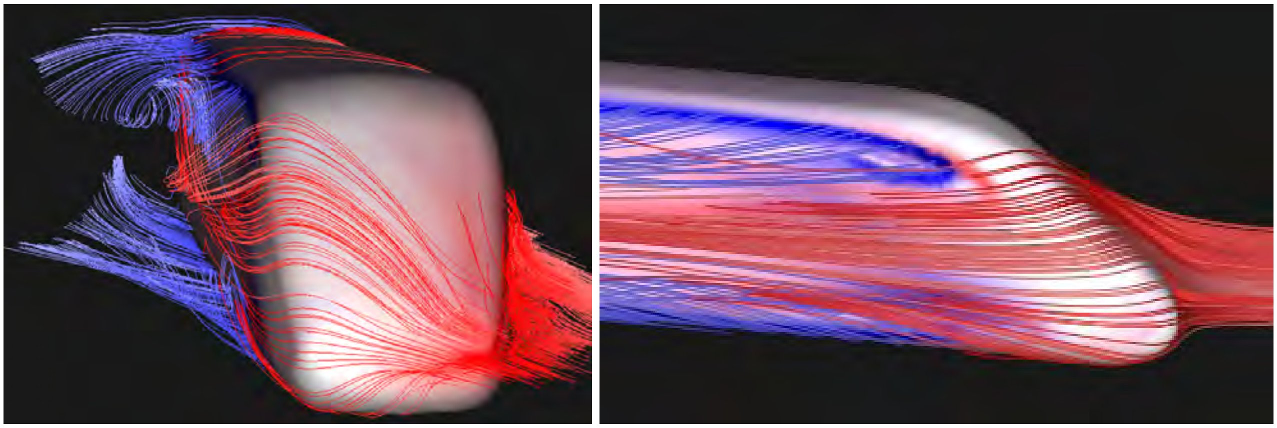


**Figure 2.** The Lagrangian visualization of the flow of air around the nose of a train reveals where sheets of air detach and reattach on the side of the train (blue corresponds to detachment, and red indicates re-attachment). This kind of analysis reveals indirect traces of large vortices forming close to the side of the train, and their study is important to determine the stability of the train at high speeds. Visualizing particle trajectories near the detachment and re-attachment zones (red and blue curves) using integral curves allows detailed insight into the flow structures.

isting methods in this context. VACET scientists have delivered visualization tools that enable fast computation and interactive visualization of vector fields using these Lagrangian techniques in its production-quality, parallel capable visual data analysis software infrastructure, VisIt.

Unfortunately, Lagrangian techniques require computing a huge number of particle traces that densely cover the domain of interest. Even for medium-sized datasets, it is not uncommon that application of Lagrangian visualization techniques requires computing millions of trace particles. In the past, VACET researchers have successfully worked to reduce the number of required particles, but the computational effort required to apply Lagrangian methods to very large, petascale datasets is still significant. This problem will be addressed by the development of novel parallelization techniques that will allow integral curve computation to scale to large supercomputers. These new capabilities, together with the Lagrangian visualization techniques that are build on them, will be implemented in production-quality, parallel-capable visual data exploration software that runs on virtually all modern platforms, ranging from desktop-class machines to DOE's petascale computer systems.

This represent a major new capability for domain scientists concerned with vector field analysis, as it will enable broad use of modern visualization methods and allow the treatment of petascale datasets when run on large parallel computing platforms.

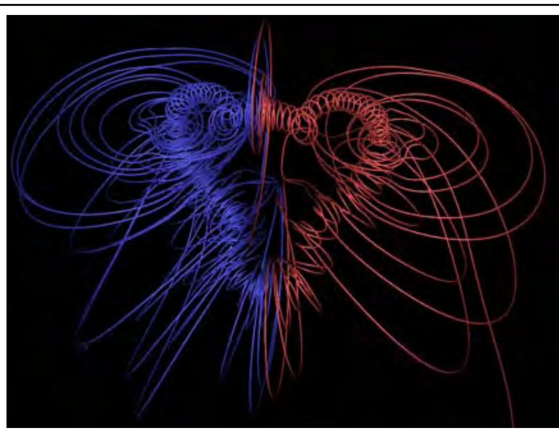


**Figure 3.** The figure depicts streamlines from an adaptive mesh refinement (AMR) computation of two incompressible vortex rings merging. In this time step, the vortex ring cores have already merged producing a complex flow field.

## Recent Publications


C. Garth, A. Wiebel, X. Tricoche, K. I. Joy, G. Scheuermann, *Lagrangian Visualization of Flow-Embedded Surface Structures*, in Computer Graphics Forum, Volume 27, Number 3, pp 1007--1014, 2008

C. Garth, H. Krishnan, X. Tricoche, T. Bobach, K. I. Joy, *Generation of Accurate Integral Surfaces in Time-Dependent Vector Fields*, in Proc. IEEE Visualization, 2008


## Awards

*People's Choice Award – SciDAC 2008.*

**For further information on this subject contact:**


Name: Kenneth I. Joy
Organization: University of California, Davis.
Email: joy@cs.ucdavis.edu
Phone: (530) 752-1077



Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# Streamlining Data Exploration and Visualization in Climate Science

E. Santos[1], H. Vo[1], T. Fogal[1], D. Williams[2], C. Doutriaux[2], and C. Silva[1]
[1]Scientific Computing and Imaging Institute, University of Utah; [2]Lawrence Livermore National Laboratory

**Summary**

The amount of data generated by global atmospheric and oceanic models is enormous. To help climate researchers make sense of all this data, we are working together with the *Earth System Grid Center for Enabling Technologies (ESG-CET)* to provide these researchers with access to data, information, models, analysis, visualization tools, and computational resources. Climate scientists will be able to efficiently perform comparative analysis and visualizations of their experiments, by leveraging a scientific workflow and provenance management system and climate and data analysis tools in a highly collaborative problem-solving environment.

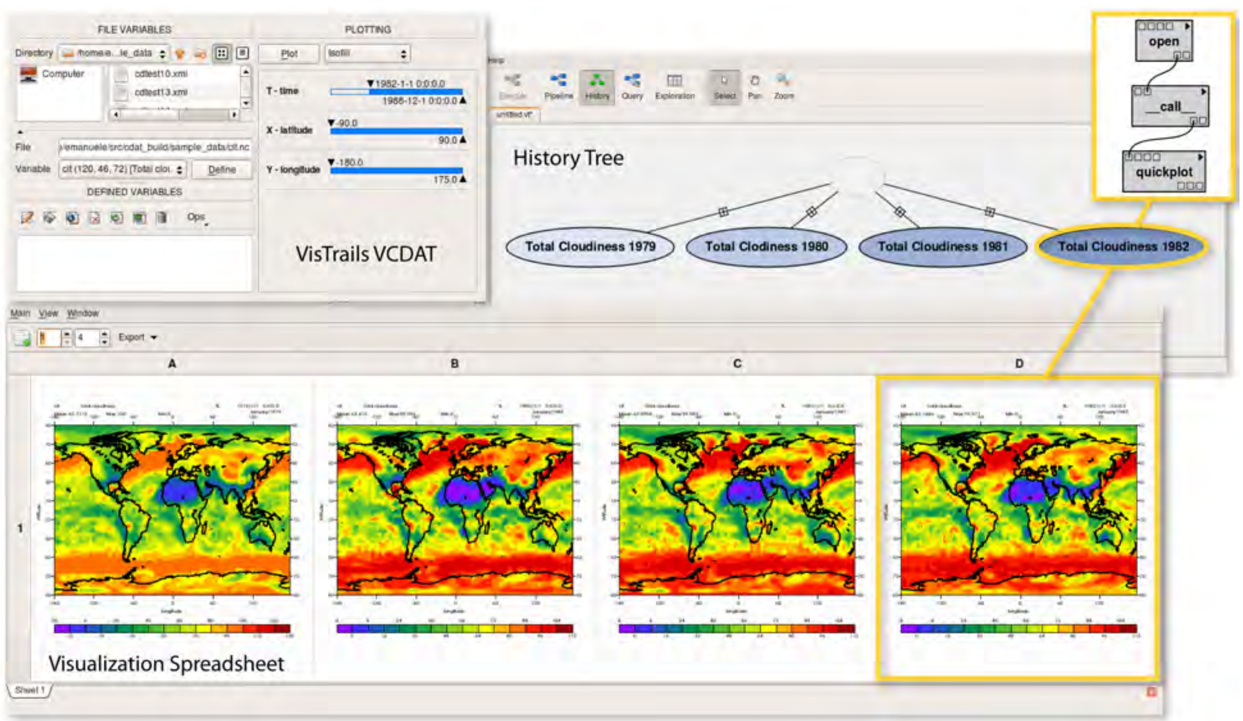



**Figure 1**. Example of comparative visualization using the CDAT VisTrails package. Using VisTrails parameter exploration mechanism and the visualization spreadsheet, a user quickly generates a side-by-side comparison of total cloudiness in the month of January from 1979 to 1982. As the user interacts with the VisTrails VCDAT window to select and manipulate variables in the climate datasets, VisTrails automatically generates and captures provenance information. This information reflects the "recipe" for creating a given visualization or series of visualizations. The "recipe" can be shared, reused, or even compared with other "recipes" to reveal differences and similarities in different types of data analysis.

One of the most difficult challenges that climate scientists face today is managing and understanding the massive amounts of global atmospheric and oceanic model data collected by observation and generated by simulations. In order to facilitate data processing, management, retrieval and analysis, the *Program for Climate Model Diagnosis and Intercomparison (PCMDI)* and other DOE sites started the *Earth System Grid (ESG)* project, which allows users anywhere to remotely access a distributed multi-petabyte archive and perform analysis of climate datasets, which are available on supercomputers and large-scale data and analysis servers. To meet its data

VACET

access and analysis needs, PCMDI also developed the *Climate Data Analysis Tools (CDAT)*, a software system that leverages the ESG infrastructure to support exploration and visualization of climate scientific datasets.

The climate community widely adopted CDAT for accessing and analyzing data in their scientific experiments, but as data size and complexity has grown, climate researchers are faced with new challenges. First, the amount of data resulting from simulations has grown too large to be easily moved across the network, making it more difficult to be analyzed and visualized. Second, scientists frequently need to compare results of simulations. For example, when they are trying to understand the effect of parameter changes on resulting simulation output. These comparisons might generate hundreds of different configurations of datasets, parameter values and visualization techniques. Keeping track of all this information using traditional (pen and paper) methods is time-consuming and error-prone.

Our team is working with the ESG-CET to address both of these challenges. The ultimate goal is to provide climate researchers worldwide with access to: data, information, models, analysis, visualization tools, and computational resources required to make sense of enormous climate simulation datasets. We started adding provenance support (see "Provenance and its Importance for Science" sidebar) and comparative visualization tools to CDAT, by integrating it to the VisTrails system. VisTrails is a scientific workflow and provenance management system developed at the University of Utah that provides support for data exploration and visualization.

The result is a CDAT VisTrails package that allows users to leverage functionality from CDAT and VisTrails, including the capability to perform comparative visualizations, while keeping the entire provenance generated during the data exploration and visualization (see Figure 1). Using a graphical interface similar to the one already available in CDAT, users can use familiar tools to build their workflows and have their entire provenance automatically captured.

***Provenance and its Importance for Science***

*Provenance* is defined by the Oxford English Dictionary as the source or origin of an object; its history and pedigree; a record of the ultimate derivation and passage of an item through its various owners. Provenance is important for science because it helps to interpret and reproduce the results of an experiment; to understand the chain of reasoning used in the production of a result; to verify that the experiment was performed according to acceptable procedures, to track who performed the experiment and who is responsible for its results.

The goal is to speed up the scientific discovery process, allowing scientists to easily examine all the steps that led to a result, identify the experiment's inputs and outputs and consequently help reproduce the results.

VisTrails brings functionality that supports concurrent exploration of multiple visualizations with the use of a spreadsheet (see Figure 1), which together with the parameter exploration mechanism allows users to effectively compare visualizations produced by different workflows side by side.

By collaborating with the PCMDI and the ESG-CET, VisTrails will be part of an overall solution to facilitate data to a worldwide audience of climate data consumers. Together, this integrated enterprise system is designed to manage and analyze extremely large and diverse datasets.

**For further information on this subject contact:**
Name: Cláudio T. Silva.
Organization: Scientific Computing and Imaging Institute, University of Utah.
Email: csilva@sci.utah.edu
Phone: (801) 587-7588

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# Visualization of Uncertainty and Ensemble Data: Exploration of Climate Modeling Data with integrated ViSUS-CDAT Systems

K. Potter[1], Y. Zheng[1], P. Bremer[2], D. Williams[2], C. Doutriaux[2], V. Pascucci[1], C. Johnson[1]
[1]Scientific Computing and Imaging Institute, University of Utah; [2]Lawrence Livermore National Laboratory

**Summary**

Climate scientists are working towards a better understanding of global climate change. Important scientific thrust areas are to better understand long-term climate impact on terrestrial life and how to mitigate these impacts and adapt to future climate conditions. To help estimate climate change, scientists perform ensemble runs, which consist of many runs of several numerical models using perturbations of input parameters and initial conditions. These ensemble runs produce massive amounts of data and give rise to the challenge to our team — to visually and analytically assist scientists in their mission to understand and estimate long-term climate change. Moreover, we are developing new capabilities for visual exploration and analysis of data collections and uncertainty information. In addition, we are creating production-quality software that is integrated into the Program for Climate Model Diagnosis and Intercomparison's (PCMDI's) globally accepted Climate Data Analysis Tools (CDAT).

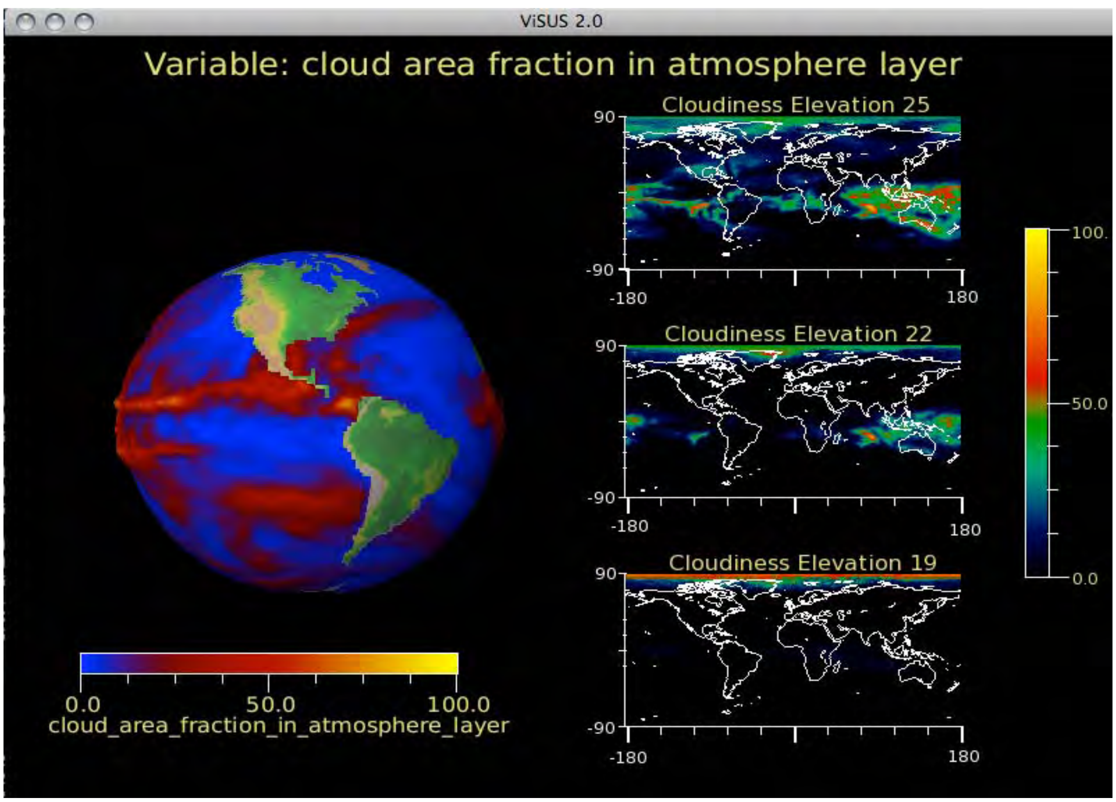



**Figure 1.** Screenshot of ViSUS 2.0 executed within CDAT. ViSUS provides the user with multiple views for data exploration and analysis. This lets the user investigate multiple variables simultaneously, as well as understand the uncertainty of a single variable. Here, the mean of the cloud area fraction is presented on the left as height on the globe and further highlighted with a redundant colormap. On the right, three elevations of cloudiness are presented using three 2D slices.

Climate scientists are generating large, complex datasets to model global changes in climate over long time periods. Such datasets combine processes underlying climate behavior such as the global carbon cycle, atmospheric chemistry, vegetation, and ocean dynamics.

In order to better understand the long-term trends in climate change, it is essential to characterize the confidence and accuracy of the simulations used for this purpose. For this reason the datasets generated by climate scientists consist of multiple simulations runs, each run using perturbed parameters and input conditions. The result is an *ensemble* dataset, which is a collection of hundreds of simulations estimating possibly hundreds of variables per grid point, across time. These datasets quickly become quite large, on the order of terabytes of data. Such datasets are therefore challenging both in data management and analysis.

Ensemble datasets, by design, provide insight into *uncertainty.* Uncertainty can be described as the accuracy or confidence associated with the data. Errors can arise in the simulation through faulty estimations of the initial conditions, finite resolution of the numerical model, and sensitivity to input parameters.

The focus of this work is to develop and enhance a production-quality data analysis and visualization tool in collaboration with the *Earth System Grid Center for Enabling Technologies (ESG-CET)*, the science application for *A Scalable and Extensible Earth System Model for Climate Change Science*, and the climate modeling community in general. The large scale and high complexity of this type of data, data management and visualization is challenging. Our solution is founded on the tight integration of 3D visualization tools, statistical data analysis techniques and compelling metaphors that facilitate the user intuition about the uncertainty of large, multivariate, time-varying datasets.

Figure 1 demonstrates a scripted visualization template generated with the ViSUS system, designed to integrate 3D visualization capabilities into the *Climate Data Analysis Tools (CDAT)* package. CDAT is specifically designed for the needs of climate scientists, providing advanced data analysis capabilities combined with the ability of reading specific data formats and providing geospatial information. The integration with ViSUS provides (among other things) a flexible system for the visualization of large ensemble datasets. Data can be displayed using a variety of techniques, including methods familiar to domain experts. 2D visualizations of the data allow for the direct comparison of multiple variables, or the variation of some dimension of a single variable. Visualizing data on the globe provides a compelling presentation, and integration with more complex 3D techniques such as iso-surfacing. Uncertainty can be displayed for example by using height and color maps, providing visual assessment of the confidence. In addition, the time component of the data can be explored by animating the datasets through each time step.

Within the CDAT framework, the VACET and ESG-CET teams are working on advancing the 2D and 3D capabilities of the ViSUS system for the advancement of climate researcher. To this end, we are working to improve visualization techniques including visual data analysis and sophisticated user interaction. We have Python scripts to facilitate the quick interchange of data sets and provide simple initial visualization settings for the user. The integrated ViSUS-CDAT system is designed to help further scientific discovery by providing a complex visualization tool in a manner directly accessible to the researchers in need.

Thus far, VACET and ESG-CET researchers have "just scratched the surface" of what is possible. Future work will include adapting more advanced visualization and full integration of the ViSUS large data management capabilities within the CDAT system for the benefit of the climate research community.

**For further information on this subject contacts:**


- Name: Kristin Potter
  Organization: SCI Institute, University of Utah
  Email: kpotter@cs.utah.edu
- Name: Dean Williams
  Organization: PCMDI, LLNL
  Email: williams13@llnl.gov
- Name: Valerio Pascucci
  Organization: School of Computing and SCI Institute, University of Utah
  Email: pascucci@acm.org


Note: This work was funded in part by the SciDAC2 VACET and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# Topological Feature Extraction for Quantitative Analysis of Terascale Combustion Simulation Data

A. Mascarenhas[2], R. W. Grout[2], P.-T. Bremer[3], E. R. Hawkes[2], V. Pascucci[3], J.H. Chen[2]
[1]Scientific Computing and Imaging Institute, University of Utah
[2]Sandia National Laboratory
[3]Lawrence Livermore National Laboratory

**Summary**

One major challenge to improving modern energy production methods is the need for better understanding of the key chemical processes that underlie the complex mechanisms of combustion. This challenge has motivated the development of advanced computational methods aimed at the Direct Numerical Simulation of the full three-dimensional dynamics of turbulent flames. Better understanding of such flames will enable new insights into reducing pollutants and increasing efficiency in combustion devices. The objective of this collaborative research is twofold: first, to develop a method for characterizing the mixing length scales on an instantaneous and local basis; second, to explore the interaction between mixing and autoignition. These goals are accomplished via a new family of robust combinatorial methods for identifying, segmenting, and tracking in time topological features related to local mixing rates and to a scalar representative of autoignition. This research could have long-term impact in a number of application areas related to energy production like aircraft engines, where fuel and oxidizers are not premixed for safety reasons, and in direct-injection internal combustion engines where diesel jet flames are stabilized downstream of the fuel injector in a hot ignitive coflow.

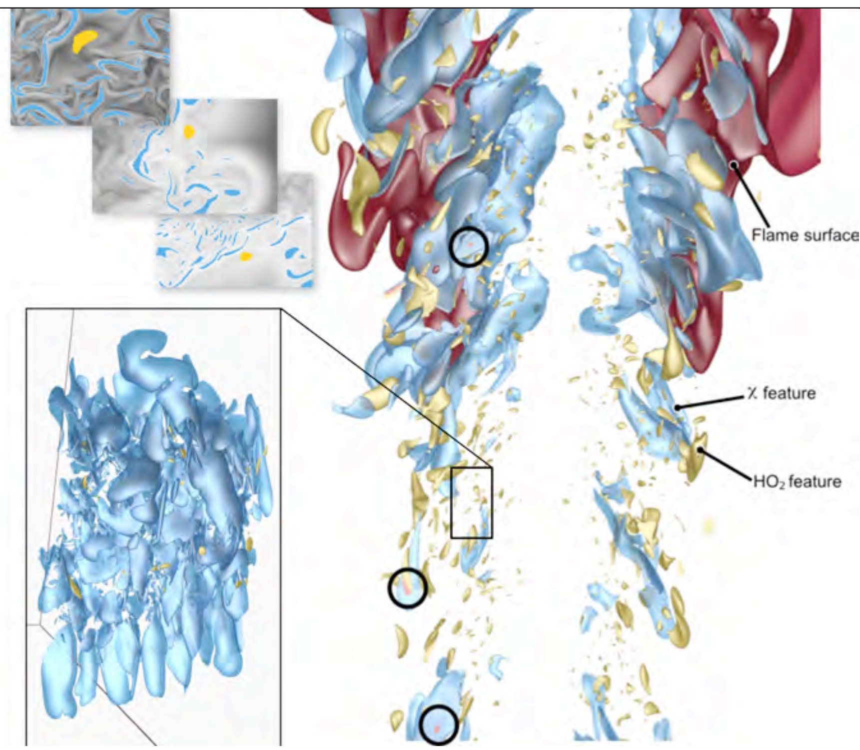



**Figure 1.** Topological analysis of a combustion simulation. Maxima of HO2 species (gold) can indicate auto-ignition kernels. These kernels are expected to live outside high-$\chi$ regions (blue regions). High-$\chi$ regions are large and many compared to $HO_2$ regions. Right, all HO2 kernels below the base of the flame with high-$\chi$ segments that overlap them. Bottom-left, close-up side view showing all HO2 kernels and high-$\chi$ segments. Top-left, three cross sections showing the relation between the HO2 and $\chi$ fields.

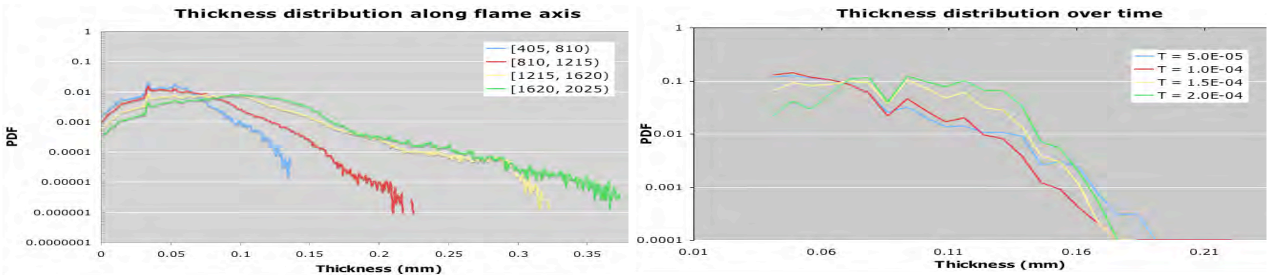



**Figure 2.** Distribution of the thickness of the high-χ features associated with extinction and reignition regions for two DNS simulations. (left) Steady case also called lifted jet. (right) Time analysis for the unsteady case.

In turbulent non-premixed combustion, where the fuel and oxidizer reactant streams are segregated, the reactant streams must be molecularly mixed before reaction can occur. Therefore, the turbulent mixing rate is a key quantity in determining the overall burning rate and efficiency. In general, as the mixing rate increases, reaction rates increase and the overall efficiency increases. A nonpremixed flame has a well-defined internal structure. Beyond a critical rate of mixing, reactions cannot keep up with the mixing and the flame quenches locally. This undesirable situation can lead to increased emissions. If quenching is pervasive, then global blow-out can occur, which would be catastrophic, for example, in an aero gas-turbine engine. In an autoignition situation, there is a similarly well-structured relationship between the species concentrations but the relationship is in time instead of space as radical concentrations build up to sufficient levels to establish a flame. Turbulent mixing is characterized locally by the scalar dissipation rate, χ, which is equal to twice the product of the molecular diffusivity and the square of the mixture fraction gradients.

3D measurements have shown that the thickness of χ scales with the small-scale turbulence, i.e. the Kolmogorov or Batchelor scales rather than with the integral scale of turbulence. The advent of terascale 3D direct numerical simulations (DNS) enabled the direct computation of χ and its evolution. The method presented here is devised to use the resolved χ field to identify and measure the structure of extinction and autoignition regions.

We study the results from two DNS simulations. The First is a temporally-evolving turbulent CO/H2 jet flame undergoing extinction and reignition at different Reynolds numbers (left of Figure 3). The second is a spatially-evolving lifted turbulent Ethylene/air jet flame. This simulation, performed on a grid of 1 billion grid points, is depicted on the right of Figure 3. In this arrangement, the configuration is statistically stationary in time.

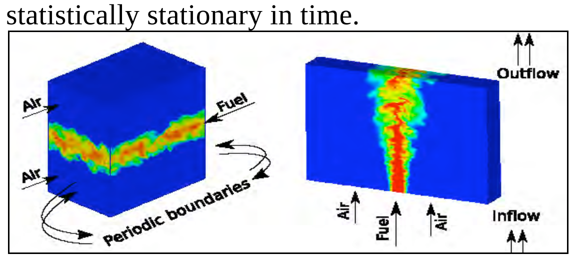



**Figure 3.** Schematic of combustion scenarios. Left, the temporally evolving jet. Right, the lifted jet.

With our topological approach we are able to provide robust segmentation and tracking of high-χ regions (Fig.1). This new capability enables the first-ever quantitative analysis of the thickness of these features and their distribution in space and time. These statistics are shown in Figure 2 for both the lifted jet case (on the left) and for the dynamic case (on the right). Our method is capable of determining a highly localized measure of the mixing length and shed new insight in the detailed dynamics of extinction and reignition processes. Application to terascale data sets indicates that the method can be applied to very large datasets on modest hardware. This is a particularly useful since the hardware used for the original simulation is out of reach for most of the combustion community and, using higher end resources, gets are closes to the petascale data analytics regime.

**For further information on this subject contacts:**

Name: Ajith Mascarenhas
Organization: CRF, Sandia National Laboratory
Email: aamasca@sandia.gov
Name: Valerio Pascucci
Organization: SCI Institute, University of Utah
Email: pascucci@acm.org

Note: This work was funded in part by the SciDAC2 VACET and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# VisIt: a Production Tool for Visualizing and Analyzing Large Data

H. Childs[1], S. Ahern[2], E. Deines[3], T. Fogal[4], C. Garth[3], J. Meredith[2], Prabhat[5], D. Pugmire[2], O. Rübel[3/5], A. Sanderson[4], G. Weber[5], B. Whitlock[1]

[1]Lawrence Livermore National Laboratory; [2]Oak Ridge National Laboratory; [3]University of California, Davis; [4]University of Utah; [5]Lawrence Berkeley National Laboratory

## Summary

The Visualization and Analytics Center for Enabling Technologies (VACET) deploys much of its research to SciDAC stakeholders via the VisIt visualization and analysis tool. VisIt, an R&D 100 winner, has two main focii: processing massive data and providing an end-user production tool. VACET has worked to further both of these goals by preparing VisIt for the petascale and by adapting VisIt to meet the needs of SciDAC users spanning many scientific areas, such as fusion, astrophysics, climate, and turbulent flow. VACET's strategy of using VisIt as a delivery vehicle for its research has several benefits: quickly and easily delivering results to stakeholders, minimizing development and deployment costs, and addressing long-term software maintenance and support concerns.


**Figure 1.** VisIt-produced SciDAC review cover images. Left: the net ecosystem exchange. Middle: collapse of a type II supernova. Right: confinement and dispersion of small molecules within carbon nanostructures.

To gain scientific insight, SciDAC scientists need to investigate the massive data sets their simulations produce. Their investigations take many forms: traditional visualization, quantitative and comparative analyses, visual debugging, as well as communication of these results. Further, these scientists need easy-to-use and richly featured tools, as the types of investigation they want to perform change from moment to moment. Further, they need tools because the number of simulations being run far exceeds the number of visualization and analysis experts available to provide assistance. VACET researchers are ensuring that SciDAC scientists can use VisIt for their day-to-day needs, and also are using it to deliver VACET's research to them.

VisIt is an open source, end user visualization and analysis tool with a strong emphasis on large data. It has a unique contract-based system that allows it to adaptively apply optimizations and also scale on large number of processors to handle the biggest data sets. VisIt has been shown to scale up to 4096 processors, and has been used to visualize data sets as big as 64 billion cells (even with only 128 processors!). Further, VisIt is undergoing Joule code certification, which will further demonstrate its scalability. VisIt follows a client-server design, with the server being

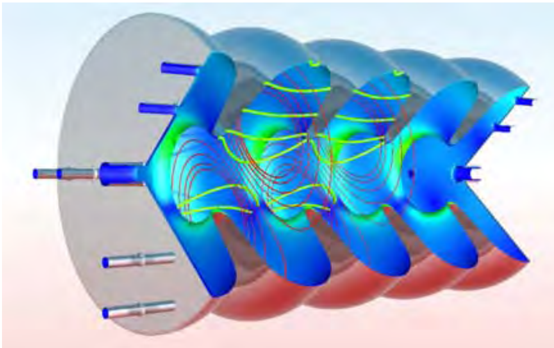


**Figure 2.** This image, created by John Cary of FACETS using VisIt, was part of a video animation that won a *People's Choice Award* at SciDAC 2008's Visualization Night.

parallelized. This architecture allows data to be visualized and analyzed where it was simulated, avoiding data movement. Further, the design allows for the client to run locally, enabling interactive graphics.

VACET has deployed VisIt to many stakeholders in the Office of Science community. Customers include many SciDAC Science Applications and Centers: Framework Application for Core-Edge Transport Simulations (FACETS); the Community Petascale Project for Accelerator Science and Simulation (COMPASS); the Computational Astrophysics Consortium (CAC); Simulation of Turbulent Flows with Strong Shocks and Density Variations (Shocks); the Global Cloud Resolving Model (GCRM); the Center for Extended Magnetohydrodynamic Modeling (CEMM); and the Applied Partial Differential Equations Center (APDEC). Further, VACET has collaborated with other SciDAC centers to deploy research in VisIt, namely the Scientific Data Management Center (SDM), and, to a lesser extent, the Interoperable Technologies for Advance Petascale Simulations Center (ITAPS), as well as with institutional visualization experts at Office of Science Laboratories, to ensure that VisIt is available and runs well on their platforms, and with SAPs for fusion and climate.

The research and software engineering performed by VACET has significantly improved VisIt. VACET transformed VisIt's streamline algorithm from one of the tool's weak spots into a strong point. VACET's integration of FastBit bitmap indexing enabled the interactive browsing of massive amounts of particles. VACET has greatly improved VisIt's volume renderer, including multi-variable support and improved shading. VACET improved VisIt's AMR infrastructure to the point that the APDEC retired their homegrown, AMR-specific visualization tool in favor of VisIt, saving them development costs. Further, VACET greatly improved the core infrastructure of VisIt. In addition to the standard bug fixes, porting changes, and small enhancements, VACET is responsible for making VisIt's software repository open to the public (instead of only available to Lawrence Livermore employees), and making a public, archived mailing list.

VisIt was originally developed by the NNSA's Advanced Simulation & Computing (ASC) program, which culminated in an R&D100 award in 2005. Its development continued and the project is now co-developed by ASCR, the NNSA, and the Office of Nuclear Energy, in addition to developers at universities, foreign laboratories, and support from private industry. The VisIt software repository has twenty-five developers from ten institutions. When VACET was originally funded, VisIt represented approximately fifty man-years of initial investment, and that investment has been further leveraged by ongoing activities outside of VACET. That said, VACET is a dominant force in VisIt development, accounting for approximately one third of all development work.

VisIt is extensively used for production visualization work around the world. It has been downloaded over one hundred thousand times and is used on supercomputers worldwide. The "visit-users" mailing list has almost three hundred subscribers, and receives approximately three hundred posts per month. SciDAC users can bypass this mailing list and directly access VisIt developers using the "visit-help-scidac" mailing list, which is supported by VACET.

**For further information on this subject contact:**

Name: Hank Childs
Organization: Lawrence Livermore National Laboratory.
Email: childs3@llnl.gov
Phone: (925) 422-4035

Note: This work was funded in part by the SciDAC2 Visualization and Analytics Center for Enabling Technologies and ASCR's Visualization Base Program by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

# 9 Publications, Presentations, Awards, Service, and Outreach

## 9.1 Publications

### 9.1.1 Peer-Reviewed Journal Articles

**2009**


1. M. Day, J. Bell, P.-T. Bremer, V. Pascucci, V. Beckner, and M. Lijewski. Turbulence effects on cellular burning structures in lean premixed hydrogen flames. *Combustion and Flame*, to appear.

2. John C. Anderson, Christoph Garth, Mark A. Duchaineau, and Kenneth I. Joy. Interactive Visualization of Function Fields by Range Space Segmentation. *Computer Graphics Forum (Proc. of Eurographics/IEEE-VGTC Symposium on Visualization 2009)*, 28(3):(to appear), 2009.

3. G. Weber, P.-T. Bremer, V. Pascucci, M. Day, and J. Bell. Feature Tracking Using Reeb Graphs. In *Proceedings of the 3rd TopoInVis Workshop*, 2009. To appear.


**2008**


1. Attila Gyulassy, Peer-Timo Bremer, Bernd Hamann, and Valerio Pascucci. A Practical Approach to Morse-Smale Complex Computation: Scalability and Generality. *IEEE Trans. Vis. Comput. Graph*, 14(6):1619–1626, 2008.

2. Janine Bennett, Valerio Pascucci, and Kenneth I. Joy. A genus oblivious approach to cross parameterization. *Computer Aided Geometric Design*, 25(8):592–606, 2008.

3. Herbert Edelsbrunner, John Harer, Ajith Mascarenhas, Valerio Pascucci, and Jack Snoeyink. Time-varying Reeb graphs for continuous space-time data. *Comput. Geom*, 41(3):149–166, 2008.

4. Luke J. Gosink, John C. Anderson, E. Wes Bethel, and Kenneth I. Joy. Query-Driven Visualization of Time-Varying Adaptive Mesh Refinement Data. *IEEE Transactions on Visualization and Computer Graphics (Special Issue: Proceedings of IEEE Visualization 2008)*, 14(6), November/December 2008. LBNL-803E.

5. Oliver Rübel, Gunther H. Weber, Min-Yu Huang, E. Wes Bethel, Mark D. Biggin, Charless. C. Fowlkes, C. Luengo Hendriks, Soile. V. E. Keränen, Michael B. Eisen, David W. Knowles, Jitendra Malik, Hans Hagen, and Bernd Hamann. Integrating data clustering and visualization for the analysis of 3d gene expression data. *IEEE Transactions on Computational Biology and Bioinformatics*, 2008. LBNL-382E.

6. C. Scheidegger, J. Schreiner, B. Duffy, H. Carr, and C. Silva. Revisiting Histograms and Isosurface Statistics. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization 2008)*, 14(6), 2008.

7. C. Dietrich, J. Comba, L. Nedel, C. Scheidegger, and C. Silva. Edge Groups: A New Approach to Understanding the Mesh Quality of Marching Methods. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization 2008)*, 14(6), 2008.

8. D. Koop, C. Scheidegger, S. Callahan, J. Freire, and C. Silva. VisComplete: Data-driven Suggestions for Visualization Systems. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization 2008)*, 14(6), 2008.

9. R. M. Kirby and C. Silva. The Need for Verifiable Visualization. *IEEE Computer Graphics and Applications*, 2008.

10. J. Daniels, C. Silva, J. Shepherd, , and E. Cohen. Quadrilateral Mesh Simplification. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2008)*, 2008.

11. J. Daniels, T. Ochotta, L. Ha, and C. Silva. Spline-Based Feature Curves from Point-Sampled Geometry. *The Visual Computer*, 24(6):449–462, 2008.

12. L. Bavoil, S. Callahan, and C. Silva. Robust Soft Shadow Mapping with Depth Peeling. *Journal of Graphics Tools*, 13(1):19–30, 2008.

13. C. Scheidegger, D. Koop, E. Santos, H. Vo, S. Callahan, J. Freire, and C. Silva. Tackling the Provenance Challenge One Layer at a Time. *Concurrency And Computation: Practice And Experience*, 20(5):473–483, 2008.

14. S. P. Callahan, J. H. Callahan, C. E. Scheidegger, and C. Silva. Direct Volume Rendering: A 3D Plotting Technique for Scientific Data. *Computing in Science and Engineering*, 10(1):88–92, 2008.

15. Gunther H. Weber, Oliver Rübel, Min-Yu Huang, Anagla H. DePace, Charless C. Fowlkes, Soile V.E. Keränen, Cristian L. Luengo Hendriks, Hans Hagen, David W. Knowles, Jitendra Malik, Mark D. Biggin, and Bernd Hamann. Visual Exploration of Three-dimensional Gene Expression using Physical Views and Linked Abstract Views. *IEEE Transactions on Computational Biology and Bioinformatics*, 2008. LBNL-6377.

16. Cecilia Aragon, Stephen Bailey, Sarah Poon, Karl Runge, and Rollin Thomas. Sunfall: A Collaborative Visual Analytics System for Astrophysics. *Journal of Physics Conference Series – SciDAC 2008*, 125, June 2008. LBNL-657E.

17. C. C. Fowlkes, C. L. Luengo Hendriks, S. V. E. Keränen, G. H. Weber, O. Rübel, M.-Y. Huang, S. Chatoor, A. H. DePace, L. Simirenko, C. Henriquez, A. Beaton, R. Weiszmann, S. Celniker, B. Hamann, D. W. Knowles, M. D. Biggin, M. B. Eisen, and J. Malik. A Quantitative Spatio-temporal Atlas of Gene Expression in the Drosophila Blastoderm. *Cell*, 133:364–374, April 2008.

18. Brian Paul, Sean Ahern, E. Wes Bethel, Eric Brugger, Rich Cook, Jamison Daniel, Ken Lewis, Jens Owen, and Dale Southard. Chromium Renderserver: Scalable and Open Remote Rendering Infrastructure. *IEEE Transactions on Visualization and Computer Graphics*, 14(3), May/June 2008. LBNL-63693.

19. John C. Anderson, Christoph Garth, Mark A. Duchaineau, and Kenneth I. Joy. Discrete Multi-Material Interface Reconstruction for Volume Fraction Data. *Computer Graphics Forum (Proc. of Eurographics/IEEE-VGTC Symposium on Visualization 2008)*, 27(3):1015–1022, 2008.

20. Christoph Garth, Xavier Trichoche, and Kenneth I. Joy. Lagrangian Visualization of Flow-Embedded Surface Structures. *Computer Graphics Forum (Proc. of Eurographics/IEEE-VGTC Symposium on Visualization 2008)*, 27(3):767–774, 2008.

21. R. Westermann K. Buerger, J. Krueger. Direct Volume Editing. *IEEE Transactions on Visualization and Computer Graphics (Special Issue: Proceedings of IEEE Visualization 2008)*, 14(6), November/December 2008.

22. M. Chen, M. Kirby, C. Hansen, and T. Moller. A Staggered Grid for Multi-resolution Volume Rendering. *Computer Graphics Forum Journal*, Accepted, to appear.

23. Aaron Knoll, Younis Hijazi, Charles Hansen, Ingo Wald, and Hans Hagen. Interactive Ray Tracing of Arbitrary Implicit Functions. *Computer Graphics Forum Journal*, Accepted, to appear.

24. Aaron Knoll, Ingo Wald, and Charles Hansen. Coherent Multiresolution Isosurface Ray Tracing. *The Visual Computer*, 2008. To Appear.

25. Guo-Shi Li, Charles Hansen, and Xavier Tricoche. Physically-based Dye Advection for Flow Visualization. *Computer Graphics Forum Journal*, 27(3):727–734, 2008.

26. Jens Bellmann, Frank Michel, Eduard Deines, Martin Hering-Bertram, Jan Mohring, and Hans Hagen. Sound Tracing: Rendering Listener Specific Acoustic Room Properties. *Computer Graphics Forum*, 27(3):943–950, 2008.

27. Scott E. Dillard, Vijay Natarajan, Gunther H. Weber, Valerio Pascucci, and Bernd Hamann. Topology-guided Tessellation of Quadratic Elements. *International Journal of Computational Geometry and Applications (IJCGA) [to appear]*, 2008. LBNL-63771, in press.

28. C. Dietrich, J. Comba, L. Nedel, C. Scheidegger, J. Schreiner, and C. Silva. Improving Mesh Quality of Marching Cubes Using Edge Transformations. *IEEE Transactions on Visualization and Computer Graphics [to appear]*, 2008.

29. Antonio Baptista, Bill Howe, Juliana Freire, David Maier, and Claudio T. Silva. Scientific Exploration in the Era of Ocean Observatories. *Computing in Science & Engineering*, 10(3):53–58, May-June 2008.

30. Erik W. Anderson, James P. Ahrens, Katrin Heitmann, Salman Habib, and Claudio T. Silva. Provenance in Comparative Analysis: A Study in Cosmology. *Computing in Science & Engineering*, 10(3):30–37, May-June 2008.

31. Juliana Freire, David Koop, Emanuele Santos, and Claudio T. Silva. Provenance for Computational Tasks: A Survey. *Computing in Science & Engineering*, 10(3):11–21, May-June 2008.

32. Claudio T. Silva and Joel E. Tohline. Computational Provenance. *Computing in Science & Engineering*, 10(3):9–10, May-June 2008.

33. Joel Daniels, Tilo Ochotta, Linh Ha, and Cláudio T. Silva. Spline-Based Feature Curves from Point-Sampled Geometry. *The Visual Computer*, 2008. To appear.

34. A.R. Sanderson, M.D. Meyer, R.M. Kirby, and C.R. Johnson. A Framework for Exploring Numerical Solutions of Advection-Reaction-Diffusion Equations Using a GPU-Based Approach. *Computing and Visualization in Science*, 2008.

35. Guo-Shi Li, Xavier Tricoche, Daniel Weiskopf, and Charles Hansen. Flow Charts: Visualization of Vector Fields on Arbitrary Surfaces. *IEEE Transactions on Visualization and Computer Graphics[accepted]*, 2008.

36. Jerry Chen, Ilmi Yoon, and E. Wes Bethel. Interactive, Internet Delivery of Visualization via Structured, Prerendered Multiresolution Imagery. *IEEE Transactions on Visualization and Computer Graphics*, Jan/Feb 2008.

**2007**

1. E. Wes Bethel, Chris Johnson, Ken Joy, Sean Ahern, Valerio Pascucci, Hank Childs, Jonathan Cohen, Mark Duchaineau, Bernd Hamann, Charles Hansen, Dan Laney, Peter Lindstrom, Jeremy Meredith, George Ostrouchov, Steven Parker, Claudio Silva, Allen Sanderson, and Xavier Tricoche. SciDAC Visualization and Analytics Center for Enabling Technology. *Journal of Physics Conference Series – SciDAC 2007*, 78, June 2007. LBNL-63542.

2. C. Garth, F. Gerhardt, X. Tricoche, and H. Hagen. Efficient Computation and Visualization of Coherent Structures in Fluid Flow Applications. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1464–1471, 2007.

3. Gunther H. Weber, Peer-Timo Bremer, and Valerio Pascucci. Topological Landscapes: A Terrain Metaphor for Scientific Data. *IEEE Transactions on Visualization and Computer Graphics (Special Issue: Proceedings of IEEE Visualization 2007)*, 13(6):1416–1423, November/December 2007. LBNL-63763.

4. C. Jones, K.-L. Ma, A. Sanderson, and L. Myers. Visual Interrogation of Gyrokinetic Particle Simulations. *Journal of Physics Conference Series – SciDAC 2007*, 78, June 2007.

5. C. Scheidegger, H. Vo, D. Koop, J. Freire, and C.T. Silva. Querying and creating visualizations by analogy. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1560–1567, 2007. Best Paper Award Winner.

6. I. Wald, H. Friedrich, A. Knoll, and C.D. Hansen. Interactive Isosurface Ray Tracing of Time-Varying Tetrahedral Volumes. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1727–1734, 2007.

7. Kenneth I. Joy, Mark Miller, Hank Childs, E. Wes Bethel, John Clyne, George Ostrouchov, and Sean Ahern. Frameworks for Visualization at the Extreme Scale. *Journal of Physics Conference Series – SciDAC 2007*, 78, 2007. LBNL-63762.

8. A. Wiebel, X. Tricoche, D. Schneider, Heike Jänicke, and Gerik Scheuermann. Generalized Streak Lines: Analysis and Visualization of Boundary Induced Vortices. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1735–1742, 2007.

9. Luke Gosink, John C. Anderson, E. Wes Bethel, and Kenneth I. Joy. Variable Interactions in Query-Driven Visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization 2007)*, 13(6):1400–1407, November/December 2007. LBNL-63524.

10. Attila Gyulassy, Vijay Natarajan, Bernd Hamann, Mark Duchaineau, Valerio Pascucci, Eduardo Bringa, and Andres Higginbotham. Topologically Clean Distance Fields. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization 2007)*, 13(6):1432–1439, November/December 2007.

11. Sung-Eui Yoon and Peter Lindstrom. Random-Accessible Compressed Triangle Meshes. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization 2007)*, 13(6):1536–1543, November/December 2007.

12. Valerio Pascucci, Giorgio Scorzelli, Peer-Timo Bremer, and Ajith Mascarenhas. Robust On-line Computation of Reeb Graphs: Simplicity and Speed. *ACM Trans. Graph.*, 26(3):58, 2007.

13. P.-T. Bremer, E. Mm Bringa, M. A. Duchaineau, A. G. Ghylassy, D. Laney, A. Mascarenhas, and V. Pascucci. Topological Feature Extraction and Tracking. *Journal of Physics Conference Series – SciDAC 2007*, 78, June 2007.

14. Hank Childs. Architectural Challenges and Solutions for Petascale Postprocessing. *Journal of Physics Conference Series – SciDAC 2007*, 78, June 2007.

15. Frank Michel, Eduard Deines, Martin Hering-Bertram, Christoph Garth, and Hans Hagen. Listener-based Analysis of Surface Importance for Acoustic Metrics. *IEEE Transactions on Visualization and Computer Graphics (Special Issue: Proceedings of IEEE Visualization 2007)*, 13(6):1680–1678, November/December 2007.

16. Michael Schlemmer, Manuel Heringer, Florian Morr, Ingrid Hotz, Martin Hering-Bertram, Christoph Garth, Wolfgang Kollmann, Bernd Hamann, and Hans Hagen. Moment invariants for the analysis of 2d flow fields. *IEEE Transactions on Visualization and Computer Graphics (Special Issue: Proceedings of IEEE Visualization 2007)*, 13(6):1743–1750, November/December 2007.

17. Attila Gyulassy and Vijay Natarajan and Bernd Hamann and Valerio Pascucci. Efficient Computation of Morse-Smale Complexes for Three-dimensional Scalar Functions. *IEEE Transactions on Visualization and Computer Graphics (Proceeding of IEEE Visualization 2007)*, 13(5), 2007.

18. Sean Ahern. Petascale Visual Data Analysis in a Production Computing Environment. *Journal of Physics, Conference Series – SciDAC 2007*, 78, June 2007.

19. Gunther Weber, Scott Dillard, Hamish Carr, Valerio Pascucci, and Bernd Hamann. Topology-Controlled Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics*, 13(2):330–341, March/April 2007.

20. Oliver G. Staadt, Vijay Natarjan, Gunther H. Weber, David F. Wiley, and B. Hamann. Interactive Processing and Visualization of Image Data for Biomedical and Life Science Applications. *BMC Cell Biology*, 8 (Suppl 1)(S10 (10 July)), July 2007,

21. Huy T. Vo, Steven P. Callahan, Peter Lindstrom, Valerio Pascucci, and Cláudio T. Silva. Streaming simplification of tetrahedral meshes. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):145–155, 2007.

**2006**

1. D. Laney, P.-T. Bremer, A. Mascarenhas, P. Miller, and V. Pascucci. Understanding Structure of the Turbulent Mixing Layer in Hydrodynamic Instabilities. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization 2006)*, 12(5):1053–1060, 2006. Best Application Paper Award Winner.

### 9.1.2 Conference Proceedings

**2009**

1. J.F. Shepherd and C.R. Johnson. Hexahedral Mesh Generation for Biomedical Models in SCIRun. In *Engineering with Computers*, volume 25(1), pages 97–114, 2009.

2. M. Callahan, M.J. Cole, J.F. Shepherd, J.G. Stinstra, and C.R. Johnson. A Meshing Pipeline for Biomedical Models. In *Engineering with Computers)*, volume 25(1), pages 115–130, 2009.

**2008**

1. Oliver Rübel, Prabhat, Kesheng Wu, Hank Childs, Jeremy Meredith, Cameron G. R. Geddes, Estelle Cormier-Michel, Sean Ahern, Gunther H. Weber, Peter Messmer, Hans Hagen, Bernd Hamann, and E. Wes Bethel. High Performance Multivariate Visual Data Exploration for Extemely Large Data. In *SuperComputing 2008 (SC08)*, Austin, Texas, USA, November 2008. LBNL-716E.

2. Daniela M. Ushizima, Oliver Rübel, Prabhat, Gunther H. Weber, E. Wes Bethel, Cecilia R. Aragon, Cameron G.R. Geddes, Estelle Cormier-Michel, and Bernd Hamann. Automated Analysis for Detecting Beams in Simulations. In *Proceedings of the Seventh International Conference on Machine Learning and Applications*, December 2008. LBNL-960E.

3. H. Wang, C. E. Scheidegger, and C. Silva. Optimal Bandwidth Selection for MLS Surfaces. In *IEEE International Conference on Shape Modeling and Applications (SMI) 2008*, 2008. Best Paper Award Winner.

4. E. Santos, L. Lins, J. P. Ahrens, J. Freire, and C. Silva. A First Study on Clustering Collections of Workflow Graphs. In *Second International Provenance and Annotation Workshop (IPAW 2008)*, 2008.

5. S. P. Callahan, J. Freire, C. E. Scheidegger, C. Silva, and Huy T. Vo. Towards Provenance-Enabling ParaView. In *Second International Provenance and Annotation Workshop (IPAW 2008)*, 2008.

6. T. Ellkvist, D. Koop, E. W. Anderson, J. Freire, and C. Silva. Using Provenance to Support Real-Time Collaborative Design of Workflows. In *Second International Provenance and Annotation Workshop (IPAW 2008)*, 2008.

7. L. Lins, D. Koop, E. W. Anderson, S. P. Callahan, E. Santos, C. E. Scheidegger, J. Freire, and C. T. Silva. Examining Statistics of Workflow Evolution Provenance: A First Study. In *Statistical and Scientific Database Management (SSDBM) 2008*, 2008.

8. C. E. Scheidegger, H. T. Vo, D. Koop, J. Freire, and C. Silva. Querying and Re-Using Workflows with VisTrails. In *ACM SIGMOD 2008*, 2008.

9. J. Freire and C. Silva. Towards Enabling Social Analysis of Scientific Data. In *CHI Social Data Analysis Workshop 2008*, 2008.

10. Cecilia Aragon, Sarah Poon, Gregory Aldering, Rollin Thomas, and Robert Quimby. Using Visual Analytics to Maintain Situational Awareness in Astrophysics. In *Proceedings of 2008 IEEE Symposium on Visual Analytics Science and Technology*. IEEE Computer Society Press, October 2008. LBNL-658E.

11. Gunther H. Weber, Vincent E. Beckner, Hank Childs, Terry J. Ligocki, Mark Miller, Brian van Straalen, and E. Wes Bethel. Visualization of Scalar Adaptive Mesh Refinement Data. *Numerical Modeling of Space Plasma Flows: Astronum-2007 (Astronomical Society of the Pacific Conference Series)*, 385:309–320, 2008. LBNL-220E.

12. David Pugmire, Hank Childs, and Sean Ahern. Parallel Analysis and Visualization on Cray Compute Node Linux. In *Cray Users Group Meeting*, Helsinki Finland, May 2008.

13. K. Potter, J. Krueger, and C.R. Johnson. Towards the Visualization of Multi-Dimensional Stochastic Distribution Data. In *Proceedings of The International Conference on Computer Graphics and Visualization (IADIS) 2008*, 2008.

14. Christoph Garth, Xavier Tricoche, Alexander Wiebel, and Ken Joy. On the role of domain-specific knowledge in the visualization of technical flows. In H. Theisel H. Hauser, S. Strassburger, editor, *Proc. Simulation and Visualization 2008 (SimVis '08)*, pages 107–120. SCS Publishing House, February 2008.

15. Christoph Garth, Alexander Wiebel, Xavier Tricoche, Kenneth I. Joy, and Gerik Scheuermann. Lagrangian Visualization of Flow-Embedded Surface Structures. In *Computer Graphics Forum (Proceedings of Eurovis 2008) [to appear]*, 2008.

16. J.F. Shepherd and C.R. Johnson. Hexahedral mesh generation constraints. In *Engineering with Computers*, volume 24(3), pages 195–213, 2008.

17. John Anderson, Luke Gosink, Mark Duchaineau, and Kenneth I. Joy. On Reconstructing Material Interfaces for Visualization. In *Proceedings of the 2008 European Visualization Conference (EuroVis) [to appear]*, 2008.

18. X. Tricoche, R. MacLeod, and C.R. Johnson. Visual Analysis of Bioelectric Fields. In *Proceedings of Visualization in Medicine and Life Sciences*. Springer-Verlag, 2008. to appear.

19. J. Krueger, K. Potter, R. MacLeod, and C.R. Johnson. Unified Volume Format: A General System For Efficient Handling Of Large Volumetric Datasets. In *Proceedings of The International Conference on Computer Graphics and Visualization (IADIS) 2008*, pages 19–26, 2008.

20. Louis Feng, Ingrid Hotz, Bernd Hamann, and Kenneth I. Joy. Anisotropic noise samples. *IEEE Transactions on Visualization and Computer Graphics)*, 14(2):342–354, 2008.

**2007**

1. Gunther H. Weber, Vincent E. Beckner, Hank Childs, Terry J. Ligocki, Mark Miller, Brian van Straalen, and E. Wes Bethel. Visualization Tools for Adaptive Mesh Refinement Data. In Werner Benger, Rene Heinzel, Wolfgang Kapferer, Wolfram Schoor, Mayank Tyagi, Shalini Venkataraman, and Gunther H. Weber, editors, *Proceedings of the 4th High End Visualization Workshop*, pages 12–25, Berlin, Germany, 2007. Lehmans Media. LBNL-62954.

2. S.G. Parker, S. Boulos, J. Bigler, and A. Robison. RTSL: a Ray Tracing Shading Language. In *In 2007 IEEE/Eurographics Symposium on Interactive Ray Tracing*, 2007. Best Paper Award Winner.

3. Janine Bennett, Valerio Pascucci, and Kenneth Joy. Genus oblivious cross parameterization: Robust topological management of intersurface maps. In *Proceedings of Pacific Graphics 2007*, pages 238–247, 2007.

4. Scott E. Dillard, Vijay Natarajan, Gunther H. Weber, Valerio Pascucci, and Bernd Hamann. Tesselletion of quadratic elements. In *ISAAC*, pages 722–731, 2006.

5. Christoph Garth, Gui-Shi Li, Xavier Tricoche, and Charles D. Hansen. Interactive Visualization of Coherent Structures in Transient Flows. March 2007.

6. John C. Anderson, Luke Gosink, Mark A. Duchaineau, and Kenneth I. Joy. Feature Identification and Extraction in Function Fields. In *EuroVis 2007*, pages 195–201, May 2007.

**2006**

1. P. Miller, P.-T. Bremer, W. Cabot, A. Cook, D. Laney, A. Mascarenhas, and V. Pascucci. Application of morse theory to analysis of rayleigh-taylor topology. In *10th International Workshop on the Physics of Compressible Turbulent Mixing*, 2006.

### 9.1.3 Invited Articles

1. C. Silva and J. Freire. Software Infrastructure for Exploratory Visualization and Data Analysis: Past, Present and Future. In *Journal of Physics: Conference Series (SciDAC 2008 Conference)*, 2008.

2. E. Wes Bethel, Chris Johnson, Charles Hansen, Claudio Silva, Steven Parker, Allen Sanderson, Lee Myers, Martin Cole, Xavier Tricoche, Sean Ahern, George Ostrouchov, Dave Pugmire, Jamison Daniel, Jeremy Meredith, Valerio Pascucci, Hank Childs, Peer-Timo Bremer, Ajith Mascarenhas, Ken Joy, Bernd Hamann, Christoph Garth, Cecilia Aragon, Gunther Weber, and Prabhat. Seeing the Unseeable. *SciDAC Review*, (8):24–33, Summer 2008. LBNL-472E.

3. Oliver Rübel, Gunther H. Weber, Min-Yu Huang, E. Wes Bethel, Soile V. E. Keränen, Charless C. Fowlkes, Cris L. Luengo Hendriks, Angela H. DePace, Lisa. Simirenko, Michael B. Eisen, Mark D. Biggin, Hans Hagen, Jitendra Malik, David W. Knowles, and Bernd Hamann. *PointCloudXplore 2: Visual Exploration of 3D Gene Expression.* GI Lecture Notes in Informatics. Gesellschaft fuer Informatik (GI), 2008. LBNL-249E.

4. E. Wes Bethel, C.R. Johnson, C. Aragon, Prabhat, O. Rübel, G. Weber, V. Pascucci, H. Childs, P.-T. Bremer, B. Whitlock, S. Ahern, J. Meredith, G. Ostrouchov, K. Joy, B. Hamann, C. Garth, M. Cole, C. Hansen, S. Parker, A. Sanderson, C.T. Silva, and X. Tricoche. DOE's SciDAC Visualization and Analytics Center for Enabling Technologies - Strategy for Petascale Visual Data Analysis Success. *CTWatch Quarterly*, 3(4), 2007.

5. Valerio Pascucci, Peer-Timo Bremer, and Ajith Mascarenhas. Topological Analysis Provides Deeper Insight into Hydrodynamic Instabilities. `http://www.sc.doe.gov/ascr/News/MonthlyNewsRoundup9-07.html`, September 2007.

6. E. Wes Bethel, Chris Johnson, Ken Joy, Sean Ahern, Valerio Pascucci, Hank Childs, Cecilia Aragon, Peer-Timo Bremer, Kathleen Bonnell, Martin Cole, Bernd Hamann, Charles Hansen, Daniel Laney, Ajith Mascarenhas, Jeremy Meredith, George Ostrouchov, Steven Parker, Prabhat, Dave Pugmire, Allen Sanderson, Claudio Silva, Xavier Tricoche, and Gunther Weber. SciDAC's Visualization and Analytics Center for Enabling Technologies. *SciDAC Review (to appear)*, 2008.

### 9.1.4 Book Chapters

1. G. Weber, P.-T. Bremer, V. Pascucci, M. Day, and J. Bell, "Feature Tracking Using Reeb Graphs", in *Proc. 3rd TopoInVis Workshop*, to appear.

2. Attila Gyulassy, Peer-Timo Bremer, Bernd Hamann and Valerio Pascucci. Practical Considerations in Morse-Smale Complex Computation, *in Proc. 3rd TopoInVis Workshop*, to appear.

3. Ajith Mascarenhas, Ray Grout, Peer-Timo Bremer, Valerio Pascucci, Evatt Hawkes and Jacqueline Chen. Topological feature extraction for comparison of length scales in terascale combustion simulation data, *in Proc. 3rd TopoInVis Workshop*, to appear.

4. Valerio Pascucci, Kree Cole-McLaughlin, and Giorgio Scorzelli "The TOPORRERY: computation and presentation of multi-resolution topology", *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*, Torsten Moller, Bernd Hamann, Robert Russell editors.

5. E. Wes Bethel, Hank Childs, Ajith Mascarenhas, Valerio Pascucci, and Prabhat. Scientific Data Managment Challenges in High Performance Visual Data Analysis. In Arie Shoshani and Doron Rotem, editors, *Scientific Data Management: Challenges, Existing Technology, and Deployment.* Chapman & Hall/CRC Press, 2008. (to appear).

6. Nameeta Shah, Scott E. Dillard, Gunther H. Weber, and Bernd Hamann. *Volume visualization of multiple alignment of large genomic DNA*. Mathematics and Visualization. Springer Verlag, December 2008.

7. Min-Yu Huang, Oliver Rübel, Gunther H. Weber, Cristian L. Luengo Hendriks, Mark D. Biggn, Hans Hagen, and Bernd Hamann. *Segmenting Gene Expression Patterns of Early-stage Drosophila Embryos*, pages 313–327. Springer Verlag, Heidelberg, Germany, 2008.

8. Y. Livnat, S. Parker, and C.R. Johnson. Fast Isosurface Extraction Methods for Large Image Data Sets. In Isaac Bankman, editor, *Handbook of Medical Imaging: Processing and Analysis*, chapter 44. Academic Press, 2nd edition edition, 2007. (to appear).

9. Steven G. Parker, Kostadin Damevski, Ayla Khan, Ashwin Swaminathan, and Chris R. Johnson. *Advanced Computational Infrastructures for Parallel/Distributed Adaptive Applications*, chapter The SCIJump Framework for Parallel and Distributed Scientific Computing. Wiley Press, 2007. to appear.

10. C.R. Johnson and X. Tricoche. Biomedical Visualization. In Pascal Verdonck, editor, *Advances in Biomedical Engineering*, pages 209–272. Elsevier, 2008.

11. P.-T. Bremer and V. Pascucci. *Topology-Based Methods In Visualization*, chapter A Practical Approach to Two-dimensional Scalar Topology. Springer Verlag, August 2007.

### 9.1.5 Edited Books

1. Topological Methods in Data Analysis and Visualization: Theory, Algorithms, and Applications Co-editors V. Pascucci, X. Tricoche, and H. Hagen. *Papers from the TopoInVis 2009 workshop*, Springer.

2. A. Pang, C. Hansen, and M. Chen, editors. *IEEE Visualization Conference and IEEE Information Visualization Conference Proceedings 2007*. IEEE Transactions on Visualization and Computer Graphics. IEEE, November 2007.

3. George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Paolo Remagnino, Ara V. Nefian, Meenakshisundaram Gopi, Valerio Pascucci, Jiri Zara, Jose Molineros, Holger Theisel, and Thomas Malzbender, editors. *Advances in Visual Computing, Second International Symposium, ISVC 2006, Lake Tahoe, NV, USA, November 6-8, 2006 Proceedings, Part I*, volume 4291 of *Lecture Notes in Computer Science*. Springer, 2006.

4. George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Paolo Remagnino, Ara V. Nefian, Meenakshisundaram Gopi, Valerio Pascucci, Jiri Zara, Jose Molineros, Holger Theisel, and Thomas Malzbender, editors. *Advances in Visual Computing, Second International Symposium, ISVC 2006 Lake Tahoe, NV, USA, November 6-8, 2006. Proceedings, Part II*, volume 4292 of *Lecture Notes in Computer Science*. Springer, 2006.

5. Chris Johnson served as Guest Co-Editor, Engineering with Computers, Special Issue on Computational Bioengineering, 2007.

6. Benger, W., Heinzel, R., Kapferer, W., Schoor, W., Tyagi, M., Venkataraman, S. and Weber, G.H., *Proceedings of the 4th High-End Visualization Workshop* (Obergurgl, Tyrol, Austria, June 18–22), Lehmanns Media, 2007.

#### 9.1.6 Posters

1. Oliver Rübel, Prabhat, Kesheng Wu, Hank Childs, Jeremy Meredith, Cameron G. R. Geddes, Estelle Cormier-Michel, Sean Ahern, Gunther H. Weber, Peter Messmer, Hans Hagen, Bernd Hamann, and E. Wes Bethel. Application of High-performance Visual Analysis Methods to Laser Wakefield Particle Acceleration Data. In *IEEE Visualization 2008*, Columbus, Ohio, USA, October 2008. (LBNL report number pending).

2. Oliver Rübel et al. Visualization and analysis of multidimensional data. In *Annual Workshop of the International Research Training Group "Visualization of Large and Unstructured Datasets*, Kaiserslautern, Germany, October 2008.

3. E. Wes Bethel, C.R. Johnson, C. Aragon, Prabhat, O. Rübel, G. Weber, V. Pascucci, H. Childs, P.-T. Bremer, A. Mascarenhas, B. Whitlock, S. Ahern, J. Meredith, G. Ostrouchov, K. Joy, B. Hamann, C. Garth, M. Cole, C. Hansen, S. Parker, A. Sanderson, C.T. Silva, and X. Tricoche. DOE SciDAC Visualization and Analytics Center for Enabling Technologies. In *2008 DOE ASCR CS PI Meeting*, Denver, CO, USA, April 2008.

4. Ray Grout, Evatt Hawkes, Jacqueline Chen, Ajith Mascarenhas, Peer-Timo Bremer, and Valerio Pascucci. Analysis of the Relationship Between High Scalar Dissipation Rate Features, Flow, and Combustion. In *WIPP 2008 Combustion Symposium [submitted]*, 2008.

5. Cecilia Aragon, Stephen Bailey, Sarah Poon, Karl Runge, and Rollin Thomas. Sunfall: A Collaborative Visual Analytics System for Astrophysics. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, Sacramento, CA, USA, November 2007. Best Poster Award Winner.

6. C. Jones, K.-L. Ma, A. Sanderson, and L. Myers. *Visual Interrogation of Gyrokinetic Particle Simulations*, 2007, Boston, MA, USA

7. Rübel, O., Weber, G.H., Huang, M.-Y., Bethel, E.W., Biggin, M.D., Fowlkes, C.C., Luengo Hendriks, C.L., Keränen, S.V.E., Eisen, M.B., Knowles, D.W., Malik, J., Hagen, H., Hamann, B., Applications of visualization and data clustering to 3d gene expression data, in: Kindlmann, G. and Linsen, L., eds., IEEE Visualization 2007 Posters, IEEE Computer Society Press, Los Alamitos, California.

8. P.-T. Bremer, *Topological Feature Extraction and Tracking*, 2007, Boston, MA, USA

9. Allen R. Sanderson, Xavier Tricoche, Christoph Garth, Scott Kruger, Carl Sovinec, Eric Held, and Joshua Breslau. Detection of Magnetic Nulls in Toroidal Geometry. In *48th Annual Meeting of the Division of Plasma Physics*, Baltimore, Maryland, 2006.

10. Allen R. Sanderson, Xavier Tricoche, Christoph Garth, Scott Kruger, Carl Sovinec, Eric Held, and Joshua Breslau. Visualizing Patterns in the Poincare Plot of a Magnetic Field. In *IEEE Visualization Conference Compendium 2006*, pages 34–35, Philidephia, Pennsylvania, 2006.

### 9.1.7 Technical Reports

1. Luke J. Gosink, Kesheng Wu, E. Wes Bethel, John D. Owens, and Kenneth I. Joy. Bin-Hash Indexing: A Parallel Method for Fast Query Processing. Technical Report LBNL-729E, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 94720, 2008.

## 9.2 Presentations

### 9.2.1 Invited Presentations

1. V. Pascucci, Fifth International Conference on Flow Dynamics, Japan, 2008.

2. V. Pascucci, Joint AMS-MAA Mathematics Meetings, Washington DC, January 6, 2009.

3. V. Pascucci, Ohio State University Computer Science and Engineering, July, 2008.

4. H. Childs, "Why Petascale Visualization Will Change the Rules", Blue Waters workshop, NCSA, Urbana-Champaign, Ill, October, 2008.

5. H. Childs, "VisIt: Visualization and Analysis Using Python", Invited presentation at SIAM Mini-symposium, Miami, FL, March, 2009.

6. Prabhat and Mark Howison. Application Case Studies with H5Part. In *HDF5 Workshop on Scalable Performance*, Berkeley, CA, USA, January 2009.

7. Gunther H. Weber. Integrating Data Analysis and Visualization. In *University College Dublin*, Dublin, Ireland, September 2008.

8. E. Wes Bethel. High Performance, Query-Driven Scientific Visualization: Finding Smaller Needles in Larger Haystacks. In *University of Tulsa Research Seminar*, Tulsa, OK, USA, September 2008.

9. E. Wes Bethel. Parallelism in Graphics and Visualization. In *University of Tulsa, CS6813 Guest Lecture*, Tulsa, OK, USA, September 2008.

10. E. Wes Bethel. Accelerating Visual Knowledge Discovery with Query-Driven Visualization. In *SIAM Conference on Imaging Science (IS08), Visualization and Analytics for Science Discovery*, San Diego, CA, USA, July 2008.

11. E. Wes Bethel. Scientific Visualization: The Modern Oscilloscope for Seeing the Unseeable. In *LBNL Summer Lecture Series*, Berkeley, CA, USA, June 2008. Also on YouTube at `http://www.youtube.com/watch?v=R4LLuEOHTtE`.

12. E. Wes Bethel. Modern Scientific Visualization is More Than Just Pretty Pictures. In *Numerical Modeling of Space Plasma Flows: Astronum-2008 (Astronomical Society of the Pacific Conference Series)*, St. John, USVI, June 2008.

13. Gunther H. Weber. Visualization and Analysis of Adaptive Mesh Refinement Data with VisIt. In *Numerical Modeling of Space Plasma Flows: Astronum-2008 (Astronomical Society of the Pacific Conference Series)*, St. John, USVI, June 2008.

14. Hank Childs. Why Petascale Visualization Will Change the Rules. In *International Conference on Computational Science 2008 (ICCS 2008)*, Krakow, Poland, June 2008. Keynote Presentation.

15. Hank Childs. Why Petascale Visualization Will Change the Rules. In *Center for Scalable Application Development Software (CScADS) Summer Workshop on Scientific Data Analysis and Visualization for Petascale Computing*, Snowbird, Utah, USA, July 2008.

16. Hank Childs. Petascale Visualization and VisIt. In *Colloquium Series for the Center for Computation and Technology (CCT) at Louisiana State University*, Baton Rouge, LA, USA, September 2008.

17. Gunther H. Weber. Visualization Tools for Adaptive Mesh Refinement Data. In *DOE Computer Graphics Forum*, Duck, NC, USA, April 2008.

18. Gunther H. Weber. Accelerating Remote Display Performance for GUI-based Applications. In *DOE Computer Graphics Forum*, Duck, NC, USA, April 2008.

19. David Pugmire, Hank Childs, and Sean Ahern. Parallel Analysis and Visualization on Cray Compute Node Linux. In *DOE Computer Graphics Forum*, Duck, NC, USA, April 2008.

20. David Pugmire, Hank Childs, and Sean Ahern. Parallel Analysis and Visualization on Cray Compute Node Linux. In *Cray Users Group Meeting*, Helsinki Finland, May 2008.

21. Charles Hansen. Multidimensional Transfer Functions and other GPU Methods. In *Exxon-Mobile*, Houston, TX, USA, August 2008.

22. Claudio Silva. Unstructured Grids and High-Quality Surface Reconstruction from Different Data Types. In *Exxon-Mobile*, Houston, TX, USA, August 2008.

23. Claudio Silva. VisTrails: Provenance and Data Exploration. In *National Biomedical Computation Resource (NBCR) Summer Institute*, San Diego, CA USA, August 2008.

24. Claudio Silva. Software Infrastructure for Exploratory Visualization and Data Analysis: Past, Present and Future. In *SciDAC 2008*, Seattle, WA USA, July 2008.

25. Charles Hansen. Interactive Texture-based Flow Visualization. In *LANL*, Los Alamos, NM, USA, August 2008.

26. Charles Hansen. CSAFE. In *University of Kaiserslautern Colloquium*, Kaiserslautern, GERMANY, May 2008.

27. George Ostrouchov. Data-Parallel Analysis and Graphics with R. In *DOE Computer Graphics Forum*, Duck, NC, USA, April 2008.

28. George Ostrouchov. Stalking the Interactive Terabyte with R: Data-Parallel Statistical Computing. In *University of Tennessee SOMS Seminar Series*, Knoxville, TN, USA, April 2008.

29. E. Wes Bethel. Visual Data Analysis and Data Exploration at the Extreme Scale. In *DOE Office of Science, Advanced Scientific Computing Research Principal Investigator Meeting*, Denver, CO, USA, April 2008.

30. Chris Johnson. Moving Beyond Pretty Pictures. In *IEEE Visualization 2007 (Capstone Speaker)*, Sacramento, CA, USA, October 2007.

31. E. Wes Bethel. Query-Driven Visualization Accelerates Scientific Insight. In *National Science Foundation (NSF) and Cyber-enabled Discovery and Innovation (CDI) Workshop*, Mathematical Sciences Research Institute, Berkeley CA, USA, October 2007.

32. E. Wes Bethel. Occam's Razor and Petascale Visual Data Analysis. In *2007 Falls Creek Falls Conference*, Nashville, TN, USA, October 2007.

33. E. Wes Bethel. Visualization, VACET and the SciDAC Compass Accelerator Project. In *Community Petascale Project for Accelerator Science and Simulation (COMPASS) All-Hands Meeting*, Fermi National Laboratory, Batavia IL, USA, September 2007.

34. Valerio Pascucci. Robust Extraction and Tracking of Topological Features in Scientific Data: State of the Art and Future Challenges. In *Workshop on Feature Extraction and Tracking (Sponsored by the Institute for Ultrascale Visualization)*, UC Davis, Davis CA, USA, August 2007.

35. Hank Childs. Architectural Challenges and Solutions for Petascale Visualization and Analysis. In *2007 Falls Creek Falls Conference*, Nashville, TN, USA, September 2007.

36. Valerio Pascucci. Robust Extraction and Tracking of Topological Features in Scientific Data: State of the Art and Future Challenges. In *Workshop on Feature Extraction and Tracking (Sponsored by the Institute for Ultrascale Visualization)*, UC Davis, Davis CA, USA, August 2007.

37. Hank Childs. VisIt: Visualization and Analysis. In *Computational Astrophysics Consortium SciDAC Center: All Hands Meeting*, Palo Alto, Ca, USA, April 2008.

38. Cecilia Aragon. SpectraVis: Visualization and Analytics. In *Computational Astrophysics Consortium SciDAC Center: All Hands Meeting*, Palo Alto, Ca, USA, April 2008.

39. Cecilia Aragon. Sunfall: Visual Analytics for Astrophysics. In *CAHSI Lecture Series, University of Texas at El Paso*, El Paso, TX, USA, April 2008.

40. Hank Childs. VisIt Overview. In *Scientific Data Management SciDAC Center: All Hands Meeting*, Seattle, Wa, USA, November 2007.

41. David Pugmire, Hank Childs, and Sean Ahern. Parallel Analysis and Visualization on Cray Compute Node Linux. In *DOE Computer Graphics Forum*, Duck, NC, USA, April 2008.

42. Hank Childs. VisIt Update. In *DOE Computer Graphics Forum*, Duck, NC, USA, April 2008.

43. David Pugmire, Hank Childs, and Sean Ahern. Parallel Analysis and Visualization on Cray Compute Node Linux. In *Cray Users Group*, Helsinki, Finland, May 2008.

44. Gunther H. Weber and Peter Nugent. Introduction into VisIt. In *UC Berkeley Astrophysics Computation Discussion Group*, Berkeley, CA, USA, February 2008.

45. Chris Johnson. Putting it All Together: Highlights of Recent National Reports on Computing. In *DOE ASC PI Meeting*, Monterey, CA, USA, February 2008.

46. Chris Johnson. Visual Computing and Imaging: Interdisciplinary Approaches. In *Center for Interdisciplinary Art and Technology*, Salt Lake City, UT, USA, February 2008.

47. Chris Johnson. Large Scale Visual Data Analysis. In *Ultrascale Visualization Workshop*, Reno, NV, USA, November 2007.

48. Luke Gosink, John C. Anderson, E. Wes Bethel, and Kenneth I. Joy. Variable Interactions in Query-Driven Visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization 2007)*, 13(6):1400–1407, November/December 2007. LBNL-63524.

49. Gunther H. Weber. Visual Data Analysis for the Berkeley Drosophila Transcription Network Project. In *DOE Computer Graphics Forum*, Peaceful Valley, CO, USA, April/May 2007.

50. V. Pascucci, *On-line computation of Reeb-graphs: simplicity and speed* Technical Paper presentation, Siggraph 2007, San Diego CA, August 2007.

51. Hank Childs. *Architectural Challenges and Solutions for Petascale Postprocessing.* SciDAC 2007, Boston MA, USA.

52. S. Ahern, *Petascale Visual Data Analysis in a Production Computing Environment*, SciDAC 2007, Boston, MA, USA.

53. C. Johnson, *SciDAC Visualization and Analytics Center for Enabling Technology*, SciDAC 2007, Boston, MA, USA.

54. Kenneth I. Joy, *Visualization and Analytics Advances for SciDAC Science,* SciDAC 2007, 2007, Boston, MA, USA.

55. G. Weber, *Visualization Tools for Adaptive Mesh Refinement Data*, presented at at the 2nd International Conference on Numerical Modeling of Space Plasma Flows ASTRONUM2007 (Paris, France, June 11-15, 2007) and the 4th High-End Visualization Workshop (University Center Obergurgl, Tyrol, Austria, June 17-22, 2007).

56. Chris Johnson, *Visualizing the Future*, UK e-Science All Hands Meeting, Edinburgh, UK, September 2008 (Plenary Speaker).

57. Chris Johnson, *Visualizing the Future*, 5th International Conference on Geographic Information Science, Park City, September 2008 (Plenary Speaker).

58. Chris Johnson, *Computing the Future*, US Library of Congress, Washington, DC, March 2009.

59. Chris Johnson, *Visual Computing*, ARUP Laboratories, Salt Lake City, October 2008.

60. C. Johnson, *Large-Scale Bioimaging and Visualization*, IEEE International Parallel and Distributed Processing Symposium (IPDPS), Long Beach, March 2007 (Keynote Speaker).

61. C. Johnson *Visualizing the Future*, Utah Computer Society, Salt Lake City, February 2007 (25th-Anniversary Keynote Presentation).

62. C. Johnson *Computational Bioimaging and Visualization: Challenges and Opportunities*, Center for Computational Molecular Biology, Brown University (Distinguished Lecture), December 2006.

63. C. Johnson *Visualizing the Future*, Second International Symposium on Visual Computing (ISVC06), Lake Tahoe, November 2006 (Conference Banquet Speaker).

64. C. Johnson *Visual Computing: Research Challenges*, Harvard University, Boston, May, 2007.

65. C. Johnson *Inverse Bioelectric Field Problems*, Inverse Days 2006, Tampere, Finland, December 2006.

66. C. Johnson *Visualizing Uncertainty*, Uncertainty Workshop, Society of Exploration Geophysicists Annual Meeting 2006, New Orleans, October 2006.

67. X. Tricoche *Characterizing the Topology of a Hamiltonian System Exhibiting Chaos: From the Standard Map to the Tokamak*, University of Kaiserslautern.

## 9.3 Tutorials

### 9.3.1 VisIt Tutorials

1. SciDAC 2008 Program Meeting, June 2008, Seattle WA. Presenters: Hank Childs and Sean Ahern. Approximately 20 attendees.

2. CScADS workshop, July 2008, Snowbird Utah. Presenters: Hank Childs and Jeremy Meredith. Approximately 20 attendees.

3. Princeton Plasma Physics Laboratory, September 2008. Presenter: Sean Ahern. Approximately 20 attendees.

### 9.3.2 VisTrails Tutorials

1. SciDAC 2008 Program Meeting, June 2008, Seattle WA. Presenters: Claudio Silva and Carlos Scheidegger. Approximately 20 attendees.

2. CScADS workshop, July 2008, Snowbird Utah. Presenter: Claudio Silva. Approximately 20 attendees.

## 9.4 Awards

1. Best Paper Award. Optimal Bandwidth Selection for MLS Surfaces. H. Wang, C. E. Scheidegger, and C. Silva. IEEE International Conference on Shape Modeling and Applications (SMI) 2008..

Figure 41: Sean Ahern delivers a VisIt tutorial to fusion researchers at the Princeton Plasma Physics Laboratory in September 2008.

2. SciDAC 2008 OASCARS. VACET researchers win three "People's Choice Awards" at Visualization Night at the SciDAC 2008 Program meeting in Seattle WA in June 2008, as well as one "Honorable Mention. These are listed and described in more detail here: `http://www.vacet.org/news/news2008.html#scidac`. See Figure 42.

3. Stakeholders win OASCARS at SciDAC 2008. VACET science stakeholders at Tech-X Corporation win an OASCAR at the SciDAC 2008 meeting using VACET software to show off their VORPAL code, which is used in high energy physics and fusion applications.

4. NVIDIA Recognized University of Utah as a CUDA Center of Excellence.

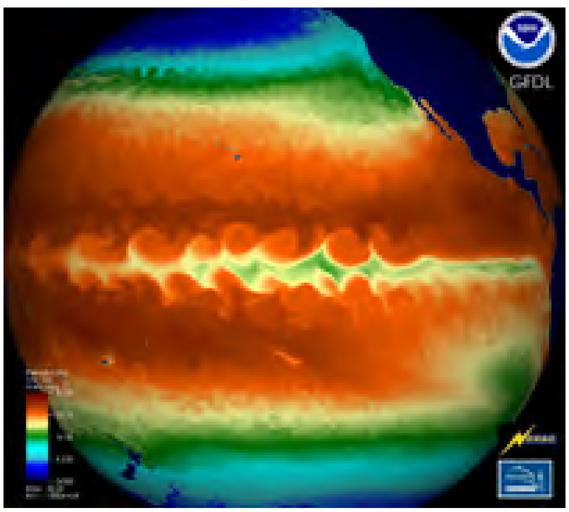
5. IEEE Visualization 2007 - Best Paper Award Winner. C. Scheidegger, H. Vo, D. Koop, J. Freire, and C.T. Silva. Querying and creating visualizations by analogy. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1560–1567, 2007. Best Paper Award Winner.

6. IEEE Visualization 2007 - Best Paper Award Nominee. Luke Gosink, John C. Anderson, E. Wes Bethel, and Kenneth I. Joy. Variable Interactions in Query-Driven Visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization 2007)*, 13(6):1400–1407, November/December 2007. LBNL-63524.

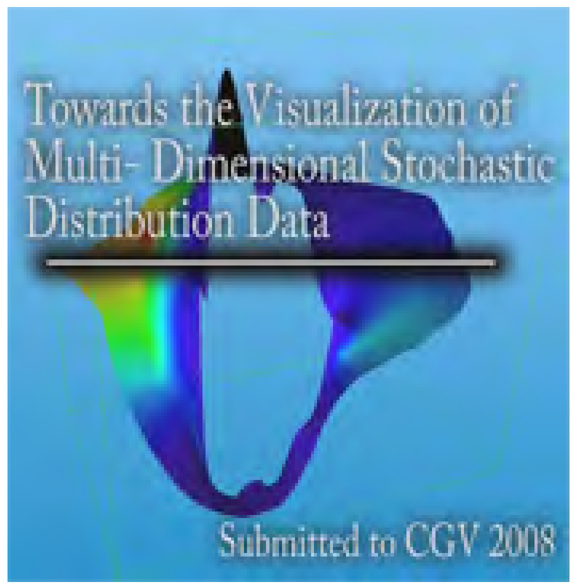
7. IEEE Visualization 2007 - Best Paper Award Nominee. C. Garth, F. Gerhardt, X. Tricoche, and H. Hagen. Efficient Computation and Visualization of Coherent Structures in Fluid Flow Applications. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1464–1471, 2007.


(a) Accurate and Efficient Integral Surfaces for Flow Visualization. C. Garth et al.

(b) The Golden Age of Scientific Visualization. C. Johnson et al.




(c) Visualization of Next-generation GFDL/NOAA Climate Simulations. Prabhat et al.




(d) Towards the Visualization of Multi-Dimensional Stochastic Distribution Data. K. Potter et al.

Figure 42: VACET researchers with three "People's Choice Awards" and one "honorable mention" at Visualization Night at the SciDAC 2008 Program meeting in Seattle, WA.

8. IEEE Raytracing Symposium 2007 - Best Paper Award Winner. S.G. Parker, S. Boulos, J. Bigler, and A. Robison. RTSL: a Ray Tracing Shading Language. In *In 2007 IEEE/Eurographics Symposium on Interactive Ray Tracing*, 2007. Best Paper Award Winner.

9. IEEE Symposium on Visual Analytics Science and Technology 2007 - Best Poster Award Winner. Cecilia Aragon, Stephen Bailey, Sarah Poon, Karl Runge, and Rollin Thomas. Sunfall: A Collaborative Visual Analytics System for Astrophysics. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, Sacramento, CA, USA, November 2007. Best Poster Award Winner.

10. IBM Faculty Award - Claudio Silva, University of Utah.

**FOR IMMEDIATE RELEASE:**

**NVIDIA RECOGNIZES UNIVERSITY OF UTAH AS A CUDA CENTER OF EXCELLENCE**

*University of Utah Latest in a Growing List of Exceptional Schools Demonstrating Pioneering Work in Parallel Computing*

**SANTA CLARA, CA & SALT LAKE CITY, UT —JULY 31, 2008**—NVIDIA Corporation, the worldwide leader in visual computing technologies, and the University of Utah today announced that the university has been recognized as a CUDA Center of Excellence, a milestone that marks the beginning of a significant partnership between the two organizations.

## 9.5 Service

### 9.5.1 Conference Chair

- Symposium on Computational Geometry 2009. General Conference co-Chair: V. Pascucci.
- 2008 IEEE Visualization, General Conference: Ken Joy, UC Davis (co-Chair).
- 2008 Volume Graphics Symposium: Valerio Pascucci, LLNL (co-Chair).
- 2008 DOE Computer Graphics Forum, Technical Program Chair: Dave Pugmire, ORNL.
- 2008 IEEE/Eurographics International Symposium on Volume Graphics. General Conference co-Chair: V. Pascucci.
- SIAM invited Minisymposium on "Visualization and Analytics for Science Discovery¡i¿ (at the SIAM Annual Meeting, 2008). Chair: V. Pascucci.
- 2007 IEEE Visualization General Conference – Ken Joy, UC Davis.
- 2007 ASCR Visualization and Analytics Workshop: Chris Johnson, Utah (co-Chair).
- 2007 Ultrascale Visualization Workshop: Chris Johnson, Utah (co-Chair).

### 9.5.2 Conference Program Committees

**2009**

- IEEE Visualization 2009: Papers Committee: Kenneth I. Joy, Christoph Garth, Valerio Pascucci, Claudio Silva
- Scidac 2009. Program Committee and Organizing Committee: V. Pascucci and H. Childs
- EUROGRAPHICS 2009. General Areas co-Chair: Jens Krüger. Program Committee: V. Pascucci.
- EuroVis 2009. Program Committee: Valerio Pascucci, Gunther H. Weber, Claudio Silva, Christoph Garth, Kenneth I. Joy, Charles Hansen

- TopoInVis 2009. Organization Committee: Valerio Pascucci; Program Committee: Peer-Timo Bremer, Gunther H. Weber, Charles Hansen, Kenneth I. Joy and Christoph Garth.
- 2009 Second ECCOMAS Conference on Computational Vision and Medical Image Processing: Chris Johnson, Utah.
- 2009 DOECGF Steering Committee: Wes Bethel and Sean Ahern. Technical Chair: Hank Childs.
- 3rd International Workshop on Knowledge Discovery from Sensor Data at KDD 2009. Program Committee: George Ostrouchov.
- Workshop on Resiliency in High Performance Computing (Resilience 2009) at HPDC. Program Committee: George Ostrouchov.
- Joint Statistical Meetings 2009. Program Chair-Elect for Section on Physical and Engineering Sciences: George Ostrouchov.

**2008**

- 2008 ACM Multimedia Technical Demonstrations: Claudio Silva.
- 2008 Knowledge-Assisted Visualization (KAV): Claudio Silva, Gunther H. Weber.
- 2008 International Symposium on Volume Graphics (VG08): Claudio Silva.
- 2008 XIX Brazilian Symp on Computer Graphics and Image Processing (SIBGRAPI): Claudio Silva.
- 2008 Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT) : Claudio Silva.
- 2008 2nd International Provenance and Annotation Workshop (IPAW): Claudio Silva.
- 2008 ACM SIGGRAPH Papers Program: Claudio Silva.
- 2008 ACM Solid and Physical Modeling Symposium (SPM): Claudio Silva.
- 2008 IEEE International Conference on Shape Modeling and Applications (SMI): Claudio Silva.
- 2008 EuroVis: Claudio Silva, Christoph Garth, Kenneth I. Joy, Charles Hansen.
- International Conference on Computer Animation and Social Agents (CASA) 2008: Claudio Silva.
- 2008 Symposium on Geometry Processing: Claudio Silva.
- 2008 SIAM Conference on Data Mining: George Ostrouchov.
- 2008 DOECGF Steering Committee: Wes Bethel and Sean Ahern.
- 2008 IEEE Workshop Knowledge-assisted Visualization: Gunther H. Weber.
- 2008 IEEE Visualization. Papers Co-Chair: Charles Hansen. Papers Committee: Kenneth I. Joy.

- 2008 EUROVIS: European Visualization Conference 2008: Kenneth I. Joy and Charles Hansen.
- 2008 Eurographics Symposium on Parallel Graphics and Visualization (EGPGV '08): Charles Hansen and Valerio Pascucci.
- 2008 Eurographics: Valerio Pascucci, LLNL.
- 2008 Symposium on Visual Computing: Valerio Pascucci, LLNL.
- 2008 Eurographics Symposium on Parallel Graphics and Visualization: Valerio Pascucci, LLNL.
- 2008 IEEE/ACM Symposium on 3D Data Processing, Visualization and Transmission: Valerio Pascucci, LLNL.
- 2008 IEEE Visualization Papers Committee: Ken Joy, UC Davis; Chuck Hansen, Utah.
- 2008 Volume Graphics Symposium: Chris Johnson, Utah; Claudio Silva, Utah
- 2008 European Visualization Conference (EuroVis): Ken Joy, UC Davis.
- 2008 Dagstuhl Visualization Conference: Ken Joy, UC Davis (co-Organizer).
- 2008 4th IEEE International Conference on e-Science: Chris Johnson, Utah.
- 2008 DOE ASCR Computer Science PI Meeting: Chris Johnson, Utah.
- 2008 International Conference on Computational Science and its Applications: Chris Johnson, Utah.
- 2008 International Society on Inverse Problems in Science and Engineering: Chris Johnson, Utah.
- 2008 VECPAR: Chris Johnson, Utah.
- 2008 IEEE International Parallel and Distributed Processing Symposium: Chris Johnson, Utah.

**2007**

- 2007 IEEE Visualization: Papers Chair, Chuck Hansen, Utah.
- 2007 IEEE Visualization Technical Program: Wes Bethel, LBNL; Valerio Pascucci, LLNL; Claudio Silva, Utah.
- 2007 IEEE Visualization Birds-of-a-Feather Chair: Gunther Weber, LBNL.
- 2007 IEEE Visualization, Local A/V Chair: Gunther Weber, LBNL.
- 2007 SciDAC Program Meeting: Wes Bethel, LBNL.
- 2007 TopoInVis: Ken Joy, UC Davis, Gunther H. Weber, LBNL.
- 2007 Conference on Interactive Techniques in Computer Graphics and Games (I3D 2007): Ken Joy, UC Davis.

- 2007 First Workshop on Knowledge-Assisted Visualization (KAV 2007): Gunther Weber, LBNL.
- 2007 International Symposium on Volume Graphics: Chris Johnson, Utah.
- 2007 IEEE International Symposium on Computer-Based Medical System: Chris Johnson, Utah.
- 2007 4th International Conference on Functional Imaging and Modeling of the Heart, Organizing Committee: Chris Johnson, Utah.
- 2007 ACM Symposium on Computational Geometry (SoCG): Valerio Pascucci.
- 2007 IEEE Conference on Shape Modeling and Applications (SMI): Valerio Pascucci.
- 2007 International Symposium on Visual Computing (ISVC): Valerio Pascucci.
- 2007 IASTED Conference on Graphics and Visualization in Engineering (VGE): Valerio Pascucci.
- 2007 IASTED Conference on Visualization, Imaging, and Image Processing (VIIP): Valerio Pascucci.
- 4th High-End Visualization Workshop (2007), Obergurgl, Austria: Gunther H. Weber
- IADIS International Conference on Computer Graphics and Visualization 2007, Lisbon, Portugal: Gunther H. Weber.

### 9.5.3 Technical Reviewer

**2009**

- EuroVis 2009. Technical Paper Reviewers: Hank Childs, LLNL, Sean Ahern, ORNL, Dave Pugmire, ORNL, Oliver Rübel, LBNL, Gunther Weber, LBNL, Christoph Gart, UC Davis Kenneth I. Joy, UC Davis, Charles Hansen, Utah..
- DOE SBIR/STTR Technical Proposal Reviewers: Wes Bethel, Prabhat, Hank Childs, Sean Ahern, Gunther Weber.
- NSF CDI-1 Pre-proposal Reviewer: George Ostrouchov.
- TopoInVis 2009. Technical Paper Reviewer: Gunther H. Weber, Oliver Rübel.

**2008**

- IEEE Transactions on Visualization and Computer Graphics. Associate Editor: V. Pascucci. Technical Paper Reviewer: Wes Bethel, Gunther H. Weber. Kenneth I. Joy, Peer-Timo Bremer, Jens Krüger, Kristi Potter, Charles Hansen, Valerio Pascucci, Christoph Garth.
- IEEE Visualization 2008. Technical Paper Reviewer: Wes Bethel, Hank Childs, Gunther H. Weber, Kenneth I. Joy, Peer-Timo Bremer, Jens Krüger, Kristi Potter, Valerio Pascucci, Christoph Garth, Oliver Rübel.
- IEEE VAST 2008. Technical paper reviewer: Gunther H. Weber.

- DOE SBIR/STTR Technical Proposal Reviewer: E. Wes Bethel.
- Journal of Systems and Software: Gunther H. Weber.
- ACM Transactions on Graphics. Technical Paper Reviewers: Peer-Timo Bremer, Jens Krüger, Valerio Pascucci.

**2007**

- IEEE Visualization 2007 Technical Program: about 12 VACET team members served as technical paper reviewers.
- IEEE Transactions on Visualization and Computer Graphics: many of the VACET team serve as technical reviewers for TVCG.
- International Journal of Software and Informatics: Gunther Weber.
- Concurrency and Computation: Practice and Experience: Gunther Weber.
- Journal of the Earth Simulator: Gunther Weber.

### 9.5.4 Program Review

- Information and Knowledge Sciences Division, Los Alamos National Laboratory: Ken Joy.

### 9.5.5 Advisory Boards and National Committees

- NSF-CRA Computing Community Consortium (CCC). Chris Johnson.
- UK Centre for Numerical Algorithms and Software for Advanced Computing - International Steering Board. Chris Johnson.
- Networking and Data Analysis Grand Challenge Center, Sandia National Laboratory, External Advisory Board. Chris Johnson.
- Computing Research Association Education Committee. Chris Johnson.
- Institute for Computational Engineering and Sciences, University of Texas, Austin, Board of Visitors. Chris Johnson.
- Finnish Centre of Excellence in Inverse Problems, Scientific Advisory Board. Chris Johnson.
- Bavarian Graduate School of Computational Engineering, International Advisory Board. Chris Johnson.
- NIH Center for Integrative Biomedical Computing, University of Utah: Ken Joy.
- Fundamental and Computational Science Directorate Review Committee, Pacific Northwest National Laboratory, (2007–). Chris Johnson.
- Scientists and Engineers for America, Advisory Board (2006–). Chris Johnson.
- Scientific Advisory Board for the Institute for Computational Engineering and Science, University of Texas, Austin, April 2007. Chris Johnson

- External Advisory Board for the NIH National Center for Microscopy and Imaging Research, UCSD, April 2007. Chris Johnson.
- External Advisory Board for the NIH National Center for Image Guided Therapy, Harvard University, June 2007. Chris Johnson.
- Scientific Advisory Board for the NIH Simbios National Center for Biomedical Computing, Stanford University, October, 2007. Chris Johnson.

#### 9.5.6 Editorial Boards

- SIAM Journal of Scientific Computing, Special Issue on Computational Science and Engineering, 2008. Chris Johnson.
- DOE Office of Advanced Scientific Computing Research Communications Project Editorial Board (2007–). Chris Johnson.
- SIAM Computational Science and Engineering, Book Series (2003–). Chris Johnson.
- Computer Graphics Year in Review (2003–). Chris Johnson.
- Electronic Transactions in Numerical Analysis (2003–). Chris Johnson.

#### 9.5.7 Panels

- *Pros and Cons of CSE Programs Versus Disciplinary Programs, the "Different Ways" to Approach Training in CSE*, SIAM Conference on Computational Science and Engineering, Miami March 2009. Chris Johnson.
- *Building a Research Group in Visualization*, IEEE Visualization 2008, Columbus, October 2008. Chris Johnson.

### 9.6 Workshops

**2009**

- Extreme Scale Workshop Series – Fusion Science Workshop, March 18-20 2009, Washington DC. Valerio Pascucci and Hank Childs.
- Extreme Scale Workshop Series – Nuclear Physics Workshop: Forefront Questions in Nuclear Science and the Role of High Performance Computing. January 26-28, 2009, Washington DC. Valerio Pascucci and E. Wes Bethel.
- 2009 TopoInVis Workshop. V. Pascucci and X. Tricoche are co-chairs of the 2009 edition of the TopoInVis workshop (see http://topoinvis.org).
- NERSC HDF5 Workshop on Scalable Performance. January 20-21, 2009. Berkeley, CA. Prabhat.

**2008**

- Extreme Scale Workshop Series – High Energy Physics Workshop: Scientific Challenges for Understanding the Quantum Universe and the Role of Computing at the Extreme Scale. December 9-11, 2008. Palo Alto, CA. E. Wes Bethel.

- Extreme Scale Workshop Series – Climate Workshop: Challenges in Climate Change Science and the Role of Computing at the Extreme Scale. November 6-7, 2008. Washington DC. E. Wes Bethel.
- Third Annual Workshop of the International Research Training Group (IRTG). Kaiserslautern, October 29 - November 1, 2008. Participants: Kenneth I. Joy, Christoph Garth, and Eduard Deines.
- Office of Nuclear Energy's NEAMS program (Nuclear Energy Advanced Modeling and Simulation) Workshop on Enabling Technologies, September 2008. Hank Childs, Wes Bethel. Childs chaired breakout session on results.
- SciDAC Framework Application for Core-Edge Transport Simulations (FACETS) Fall 2008 Program meeting. September 4-5, 2008. Dave Pugmire, Allen Sanderson.
- Center for Scalable Application Development Software (CScADS) workshop on Scientific Data Analysis and Visualization for Petascale Computing, July 2008. Hank Childs, Jeremy Meredith.
- DoE Mathematics for the Analysis of Petascale Data Workshop (MAPD), June 3-5, 2008. George Ostrouchov, Kenneth I. Joy and Valerio Pascucci.
- DOE Modeling and Simulation at the Exascale for Energy, Ecological Sustainability and Global Security (E3SGS) Town Hall Meetings. About ten different persons from VACET contributed to this effort.

**2007**

- SC2007 Workshop. Joint with the Visualization Institute, we organized a workshop at SC2007.
- SciDAC 2007 Program Meeting – Tutorial/Workshop. VACET conducted a one-day workshop in conjunction with the SciDAC 2007 Program meeting in late June 2007 in Boston, MA.
- VACET team members served in a leadership role for the 2007 prestigious Dagstuhl Seminar on "Scientific Visualization," by invitation only. K. Joy was a part of the organizing committee and C. Johnson, C. Hansen, V. Pascucci, C. Silva, were invited to give presentations.
- International Summer School on Scientific Visualization 2007. C. Hansen, and V. Pascucci lectured for a week at the summer school "Ecoles D'Ete" on the topic of "Advanced Methods in Scientific Visualization" 2007.

**2006**

- Multiple VACET presentations at the SC06 Workshop on Ultrascale Visualization: *Query-Driven Visualization Accelerates Scientific Insight* by E. W. Bethel; *VisIt Overview* by H. Childs; *Robust Topology-Based Analysis of Large Scale Data* by V. Pascucci.

## 9.7 VACET Website

VACET expends a great deal of effort to keep its website, `http://www.vacet.org`, up to date. This period, we have:

- Reorganized our Images/Movies gallery pages, and added a substantial amount of content.
- Added news items as they occur.
- Updated our publications page, as well as maintain a BibTex file of VACET publications.
- Maintained a "Software" page where visitors can download the latest version of our software applications.

## 9.8 Outreach

- Congressional outreach. U. S. Congressman Jim Matheson (Utah), who sits on the Science and Energy committee in the U. S. House of Representatives, visited SCI on 4/11/2008. As part of that visit, he dropped in on the VACET All Hands Meeting, held 4/10 and 4/11/2008 at the Scientific Computing Institute at the University of Utah. VACET members explained to Congressman Matheson the important role played by SciDAC in helping to maintain leadership in scientific research, and the crucial role that VACET plays in terms of providing production-quality software infrastructure for gaining scientific insight from large and complex scientific data being produced by simulations and collected from experiments.

Figure 43: Pictured here are VACET team members and U. S. Congressman Jim Matheson at the Scientific Computing Institute, University of Utah, 4/11/2008. Left to right: V. Pascucci (LLNL), S. Ahern (ORNL), W. Bethel (LBNL), Congressman Matheson, C. Johnson (Utah), K. Joy (UC Davis).

- HPCwire interview[12] with VACET personnel.

[12] http://www.lbl.gov/CS/Archive/news110306.html

## 9.9 Journal Covers

### 9.9.1 SciDAC Review – Spring 2009


Figure 44: Molecular dynamics simulation of confinement and dispersion of small molecules within carbon nanostructures, mimicking the dynamics of electrolytes in porous carbon materials. The visualization was generated by the SciDAC code VisIt. Simulation: Dr. Vincent Meunier, ORNL. Visualization: J. Meredith and S. Ahern, ORNL.

### 9.9.2 SciDAC Review – Special Issue 2009


Figure 45: Fluid velocity streamlines during a type II supernova collapse. Visualization by D. Pugmire (ORNL), and simulation by E. Endeve, C. Cardall, R. Budiardja (ORNL and U. Tennessee-Knoxville) and A. Mezzacappa (ORNL).

### 9.9.3 SciDAC Review – Spring 2008


Figure 46: The instantaneous Net Ecosystem Exchange (NEE) is a measure of the atmospheric-terrestrial exchange of carbon as simulated with a version of the CCSM3 model. This simulation was conducted as part of the C-LAMP project, an intercomparison activity designed to evaluate the performance of biogeochemical models when coupled to atmospheric general circulation models.

### 9.9.4 SciDAC Review – Spring 2007

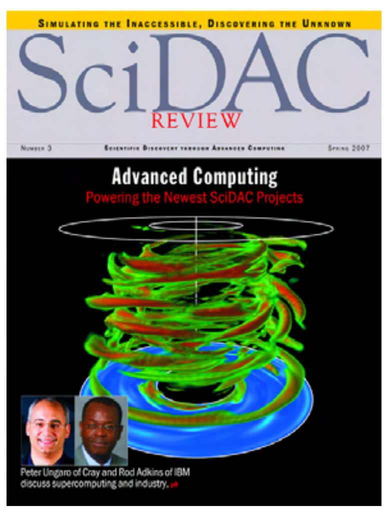



Figure 47: The SciDAC Visualization and Analytics Center for Enabling Technologies (VACET) worked with Dr. Fausto Cattaneo to visualize the time evolution of the total advective radial flux of axial angular momentum. These simulation results – carried out under DOE's INCITE program – are relevant to star and black hole formation.

# 10 Resources

## 10.1 NERSC/LBNL

VACET applied for and received an ERCAP allocation at NERSC for AY08. We were awarded 50,000 CPU hours, approximately 50TB of archival storage, and approximately 1TB on the NERSC global filesystem (GFS). Recently, we requested and received a temporary GFS quota increase to 5TB in order to support "hero-sized" visual data analysis projects in support of our science stakeholders in accelerator and combustion. That effort, conducted jointly with our stakeholders, produced novel research results as well as multiple paper submissions to IEEE Transactions on Visualization and Computer Graphics and Supercomputing 2008.

During the course of this AY, we ended up consuming way more cycles than originally forecast. The bulk of this overrun resulted from the cost of doing parallel I/O performance for the Climate/Randall project (Section 3.5). The ERCAP allocation manager, Francesca Verdier, made several additions of cycles to our allocation over the course of the project – we are grateful for her help.

We will be completing a new ERCAP allocation for resources at NERSC. The application is due 10/1/2008.

## 10.2 LCF/ORNL

VACET applied for and received a Director's Discretionary allocation at ORNL for 2008-2009. We requested and were awarded 30,000 hours on the Jaguar supercomputer, a Cray XT4/5 with 7,832 quad-core Opteron processors. This allocation includes access to the large parallel filesystem, the HPSS archival storage system, and access to the analysis and visualization cluster named lens. This allocation is shared between VACET and the SciDAC Ultravis Institute.

We will be requesting a new Director's Discretionary Allocation at ORNL when the current one expires.

## 11 Community Testimonials

COMPUTATIONAL RESEARCH DIVISION

March 24, 2009

Dr. Edward W. Bethel
50F-1603
Lawrence Berkeley National Laboratory
1 Cyclotron Road
Berkeley, CA 94720

Dear Wes,

In my capacity as the Coordinating Principal Investigator of the SciDAC Applied Partial Differential Equation Center, I am writing to express my strong support for the SciDAC Visualization and Analytics Center for Enabling Technology in their upcoming program review, and our intention of continuing to collaborate.

Prior to working with VACET, our team developed and maintained an in-house adaptive mesh refinement (AMR) visual data analysis software application (ChomboVis). VACET helped us to transition from ChomboVis to VisIt, a production-quality, parallel capable visual data analysis application. In order for us to make this transition, we worked closely with the VACET team to specify dozens of capabilities required by our applications; the VACET team then implemented these capabilities, which are now part of the production VisIt release and as such are available to the larger computational science community.

In addition to a direct cost savings from not having to support and maintain such software, our team is in a better position now to perform visual data analysis and knowledge discovery of complex, time-varying AMR datasets now and in the future as we move into the petascale regime of computing and data production. Having this kind of capability is crucial for the success of our project and for our science stakeholders, who also have either adopted or are transitioning to day-to-day use of VisIt.

Sincerely,

Phillip Colella
Senior Staff Scientist



Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road, MS: 50A-1148 | Berkeley, CA 94720 | Tel: 510•486•5412 | Fax: 510•486•6900

From:
Cameron Geddes
Staff Scientist, LOASIS program of LBNL
MS71-259, 1 Cyclotron Rd
Berkeley CA 94720, USA

To Whom It May Concern:

I am pleased to offer my support for the SciDAC Visualization and Analytics Center for Enabling Technology in their upcoming program review. I am writing as the Principal Investigator for M558, a simulation project of two million CPU-hours/year. This project supports experiments at the LOASIS program of LBNL, headed by Wim Leemans, which are developing high gradient laser wakefield particle accelerators to extend the reach of high energy physics and light sources. Simulations using the parallel VORPAL framework are used to interpret physics of the experiments (including the first narrow energy spread and first GeV beams) and to design next generation facilities such as the BELLA PetaWatt laser. The simulations generate multi-TB output that requires sophisticated visual analysis. Our team participates in the advanced accelerator component of the Community Petascale Project for Accelerator Science and Simulation (COMPASS), and received a 2006 INCITE award.

Through a productive collaborative effort with experimental, theoretical and computational physicists from LBNL and Tech-X Corp, the VACET Center has had a very positive impact on our laser wakefield simulations. Accomplishments include: developing a process for finding and tracking beam particles that requires seconds or minutes on our largest (multi-TB) simulation output instead of hours, and joint research and development of supervised and unsupervised machine learning techniques to help us quickly find and analyze beam particles in large simulation output. VACET has also provided us with production-quality, parallel visual analysis software (VisIt) that we use for exploration of simulation output too large for other tools as well as for creating high quality images. Several such images will appear in an upcoming issue of SciDAC Review. This work has resulted in joint publications in venues like the annual Supercomputing conference, the International Conference on Machine Learning Applications, and the SciDAC Review (submitted), and is now being integrated into our workflow and movie creation.

By allowing efficient analysis of large datasets, the VACET team has provided valuable support helping our team make progress towards petascale science, and is providing a valuable service to the scientific community.

Sincerely,

Dr. Cameron Geddes

***LOASIS Program***
Accelerator and Fusion Research Division
1 Cyclotron Road
Berkeley, California 94720

Mailstop 71-259
phone: (510) 495-2923
fax: (510) 486-7981
e-mail: cgrgeddes@lbl.gov

March 1, 2009

To: whom it may concern

I am writing in strong support of the VACET Center for Enabling Technology for their upcoming review. The VACET team has done an extraordinary job at making scientific visualization software, the VisIt package to be specific, not only available to, but also usable by the scientific community

I am lead PI of the FACETS SciDAC as well as an institutional PI for one of the primary institutions for the COMPASS SciDAC. We were pleased to receive Hank Childs at Tech-X in 2007 to hear about the capabilities of VisIt. It seemed probable at the time that VisIt could meet our needs, as we were in need of a scientific visualization package that would allow us to remotely visualization the large datasets that we were generating. Further, the cross-platform nature of VisIt meant that our scientists, who work from desktops and laptops running Linux, Windows, and OS X, would all be able to use VisIt. Finally, the fact that VisIt could work with ssh tunneling meant that we would be able to circumvent the firewall problems that had plagued us over the years with other solutions.

Two years have passed, and I am glad that we decided to invest in VisIt. The VisIt team helped us early on in use of VisIt and in getting our data into VisIt by giving us with advice and assistance in writing a plugin. They were responsive to our needs. We were sufficiently excited about this that we went on to develop a generic schema (VizSchema) for describing data to be visualized and modified our plugin to be compliant. Gunther Weber from the VisIt team helped us to make our plugin parallel, so that we could visualization the 20-100 GB data files that we now routinely generate in our simulations. Moreover, we are able to visualize data from both of the above SciDACs (VORPAL and FACETS data), as well as data from other SciDACs (NIMROD data for CEMM, VORPAL data for CSWIP), and even data from the Climate community (using the MODAVE project) and for materials science (the PolySwift++ computational application). Images of data from these various communities (generated by VisIt with data imported by the VizSchema plugin) are visible at https://www.txcorp.com/trac/vizschema. These visualizations have appeared on the cover and interiors of the SciDAC Review, as well. At every stage of our work, the VisIt team has been responsive to our needs and helped us meet our goals.

Perhaps of special mention is that we were able to use the VisIt application to win an OASCR People's Choice award at the 2008 SciDAC meeting in Seattle. (See https://hpcrd.lbl.gov/SciDAC08/files/vis-night.html.). The fact that VisIt works so well with other software was instrumental here, as we were able to augment the visualization by taking VisIt output and adding some special effects to it through importation into a ray-tracing program.

Given our success in using VisIt, we are now moving ahead with plans for customizing it to meet our needs for individual projects. VisIt is a powerful tool; it can nearly anything

that a user in any domain might need (and the VisIt team is working hard, I know, to provide even more capability, such as improvements to streamlines). Our intention is to build some “dumbed-down” VisIt-like applications (so called *Skins*) that would do only what is needed for a particular area of science. We have plans to do this in both fusion (for FACETS data) and in accelerators. This would be based on VisIt technology, but with choices more limited to those that make sense to our particular application areas. This will probably help widen the usage of VisIt technology in our application areas.

I expect this to succeed, in part because it is my feeling that in many respects we are not just recipients of VACET software, but we are collaborating with them and so benefit from their knowledge as well. This was and is eminently true in terms of plugin development. I have no doubt that it will continue in the development of the various Skins that we will undertake with their help.

In summary, I am sold on VisIt and the VACET team. They are providing a great service to our community in not only providing a tool, but also in providing usable tool that helps scientist extract knowledge from the huge amounts of data that are being generated by SciDAC applications now in use.

Sincerely,

John R. Cary
CEO, Tech-X
Prof. Physics, University of Colorado

Pacific Northwest
NATIONAL LABORATORY

Tel: (509) 375-6525
Fax: (509) 372-4720
Karen.schuchardt@pnl.gov

March 13, 2009

To Whom It May Concern:

I'm writing this letter to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technology in their upcoming program review. In support of my science efforts, the team has provided invaluable support helping me and my teams make progress towards petascale science by developing visualization capabilities for high resolution, geodesic grid-based data and providing an IO API and visualization capability for our new hybrid subsurface model.

I am writing as the Principal Investigator of two SciDAC Science Application Partnerships (SAP). The first SAP is providing IO and data services for Professor Dave Randall's Global Cloud Resolving Model (GCRM). The second SAP is providing a subsurface modeling user environment for Dr. Timothy Scheibe's Hybrid Numerical Model. The latter project is supported by an additional SAP, led by Dr. Bruce Palmer, which is developing the computational framework and particle-based portion of the hybrid model. In my roles, I am responsible for implementing a data strategy for each project and integrating visualization and analysis solutions.

Through a productive collaborative effort with our teams, the VACET Center has had a positive impact on our (science) projects. These accomplishments include:

- visualization capabilities developed in VisIt for handling GCRM data
- adopting H5Part as a data model and API for persisting particle-based outputs
- using VisIt for visualization of large particle datasets
- close collaboration to verify and test scalable data model for GCRM
- assistance in troubleshooting and profiling collective IO for GCRM code

Both models are radically new, ground-up development efforts. The visualization tools are invaluable to the build and test process and will be equally valuable for model analysis. In particular, the capability for remote parallel visualization is critical for these high resolution models. The ready access to the H5Part tools saved substantial time in developing a data model for the particle model and we expect to substantially reduce

the effort required completely parallelize IO in these simulations. In combination with the expert help on IO profiling and troubleshooting on Franklin, VACET has had a positive impact on our science efforts, and is providing a valuable service to the scientific community.

Sincerely,

Karen Schuchardt
Chief Scientist
Applied Computer Sciences

UNIVERSITY OF CALIFORNIA, SANTA CRUZ

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO    SANTA BARBARA • SANTA CRUZ

DEPARTMENT OF ASTRONOMY AND ASTROPHYSICS
UCO/LICK OBSERVATORY

SANTA CRUZ, CALIFORNIA 95064
FAX (831)-459-5265

woosley@ucolick.org • (831)-459-2976

March 13, 2009

Dear Reviewer:

As Principal Investigator for the SciDAC Computational Astrophysics Consortium, I am writing to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technology, especially in their upcoming program review. Their team has been essential to our progress towards petascale science. VACET's codes and expertise have allowed us to visualize the complex data sets generated in our multi-dimensional simulations of stellar explosions. They have also helped us to develop machine intelligence that will ultimately automate and optimize the comparison of large observational data sets with our models.

Of particular importance has been their assistance in transitioning to the visual data analysis software (VisIt) necessary for analyzing production-quality runs with our new AMR codes, CASTRO and MAESTRO, as well as legacy simulation codes like VULCAN and SNe. VACET has also helped our team solve challenging problems in spectral data analysis resulting from the flood of observed spectra and light curves. They have allowed us to replace "chi-by-eye" spectral fitting procedures with new objective techniques that include machine learning methods and advanced comparative visual analysis. Overall, VACET has had a positive impact on our science effort, and is providing a valuable service to the scientific community.

Yours sincerely,

Stan Woosley

S. E. Woosley
Professor of Astronomy and Astrophysics

**Princeton University**
**Princeton Plasma Physics Laboratory**
James Forrestal Campus
P.O. Box 451 MS-27, Princeton, NJ 08543-0451
(609) 243-2635

March 20, 2009

To Whom It May Concern:

I'm writing this letter to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) in their upcoming program review. I am writing as the Principal Investigator for the Center for Extended Magnetohydrodynamic Modeling (CEMM). In this role, our team is responsible for understanding the mechanisms that lead to disruptive and other stability limits in burning plasma experiments such as ITER.

The VACET Center has had a positive impact on our project by supporting our team as we move towards utilizing VisIt as our primary visualization and analysis tool. Their ability to provide tutorials on using VisIt not only at meetings such as the SciDAC PI meeting but also on site here at PPPL have been invaluable in allowing us to effectively utilize VisIt.

Their ongoing work on tools for topological analysis will help us better identify instabilities in the plasma flow which is critical to understanding the nature of tokamak experiment design. This is especially true as we move towards petascale computing and need to rely on tools that can automatically identify these instabilities. The quality of the initial analysis tool has generated interest not only within our group but within other fusion SciDACs as well that are headed by Don Batchelor and Bill Nevins . It has even been featured in a book, The Plasma Universe, by Amitava Bhattacharjee.

In addition, they have been collaborating with us to deploy tools to enable comparative analysis between our two simulation codes, M3D and NIMROD. When deployed, these tools will help us with our verification and validation efforts which are critical to the success of our project.

In summary, VACET is a valuable asset to the scientific community and we look forward to our continued collaboration.

Sincerely,

Stephen C. Jardin
Director, SciDAC Center for Extended Magnetohydrodynamic Modeling (CEMM)
Principal Research Physicist and co-Head of Computational Physics Division, PPPL
Lecturer with Rank of Professor, Astrophysics Department, Princeton University

Sandia National Laboratories

Operated for the U.S. Department of Energy's
National Nuclear Security Administration
by **Sandia Corporation**

**Jacqueline H. Chen**
Distinguished Member of Technical Staff
Reacting Flows Department (08351)

P.O. Box 969 Mail Stop 9051
Livermore, California 94551-0969

Phone: (925)294-2586
Fax: (925)294-2595
Internet: jhchen@sandia.gov

March 21, 2009

Jacqueline Chen
Combustion Research Facility
Sandia National Laboratory
Livermore, CA 94550

To Whom It May Concern:

It is a great pleasure to write this letter of support for the Visualization and Analytics Center for Enabling Technology (VACET) which has positively impacted our science productivity by enabling us to glean physical insight from terabytes of combustion simulation data.

I am writing as the lead PI of several DOE INCITE awards and as the PI of both a DOE Basic Energy Sciences Chemical Physics project on turbulent reactive flows and a DOE ASCR project on data discovery. In this role, I am leading a team of researchers responsible for the development of advanced computational methods aimed at the Direct Numerical Simulation of the full three-dimensional dynamics of turbulent flames undergoing strong finite-rate chemical processes, including local extinction and autoignition. These simulations are the first-ever 3D direct numerical simulations of fully developed turbulent flames with detailed chemical kinetics, providing a wealth of data for both fundamental understanding and model validation.

During our collaboration, VACET researchers have been collaborating closely with us to develop new advanced topological methods for the robust characterization and segmentation of extinction and reigniting regions that are generated intermittently during combustion. In a second related project VACET researchers are helping us understand the role of transient autoignition in the stabilization of lifted turbulent jet flames by exploring the use of Jacobi sets together with in-situ particle tracking in the time tracking of shortlived, small spatial ignition features. These methods have enabled us, for the first time, to robustly track over time large numbers of extinction or ignition regions, which is essential to the development of a better understanding of the flame extinction/reignition and stabilization dynamics. Better understanding of the details of such flames enable new insights into reducing pollutants and increasing efficiency in combustion devices. This research could have long-term applications in such areas as jet aircraft engines, where fuel and oxidizers are not premixed for safety reasons, and in direct-injection internal combustion engines where diesel jet flames are stabilized downstream of the fuel injector in a hot ignitive coflow.

Overall, our collaboration with the VACET team has proven to be extremely valuable and is accelerating our progress towards the understanding of combustion science from 100's of terabytes of complex, multi-scale data generated from scientific simulations. We are planning to expand our collaborative efforts with the VACET team in the coming year, and we expect that

the innovative data analysis techniques developed in the process will be extremely valuable, not only for us, but for a wide range of scientific applications.

Sincerely yours,

Jacqueline H. Chen
Distinguished Member of Technical Staff
Sandia National Laboratories

Earth System Grid

March 22, 2009

To: Scientific Discovery through Advanced Computing (SciDAC)
Office of Science
U.S. Department of Energy

From: Dean N. Williams
Lawrence Livermore National Laboratory
Program for Climate Model Diagnosis and Intercomparison
Mail Stop: L-103
7000 East Ave.,
Livermore, CA 94550

Re: Letter of Support for "Visualization and Analytics Center for Enabling Technologies (VACET)"

---

To Whom It May Concern:

As the PI of the SciDAC Earth System Grid Center for Enabling Technologies (ESG-CET), I am writing this letter to express my enthusiastic support for the activities concerning the "Visualization and Analytics Center for Enabling Technologies (VACET)". The work that the VACET team is undertaking directly benefits the ESG-CET project as well as other DOE funded programs – as it addresses some of the most challenging data visualization and analytics issues facing the climate community.

In my role as PI, I am responsible for a large multi-institutional team focused on the mission of providing climate researchers worldwide with access to distributed data, information, models, analysis tools, and computational resources required to make sense of enormous climate simulation datasets. From the start of our project, our team has been collaborating with VACET researchers to add new functionalities to our data management and visualization infrastructure. This collaboration has allowed us to introduce a completely new family of 3D visualization features that are based on ViSUS and respect our Python scripting infrastructure. In addition to the new visualization features, the ViSUS infrastructure allows us to enable external memory data processing capabilities with which we can explore interactively large datasets. Currently, we are in the process of deploying the ViSUS extensions as part of the Climate Data Analysis Tools (CDAT) 5.1 release.

The collaboration with the VACET team is also providing us with a new tool, called VisTrails, which captures full provenance of the data analysis and visualization process. This new technology will be extremely useful to climate scientists since it will allow them to share not only the data generated by their simulations but also the scripts representing the sequence of processing steps encoding their data analyses.

In conclusion, our collaboration with the researchers of the VACET team has been very useful and we are committed to continue this partnership in the future. The tools provided by the VACET team will prove very useful for climate scientists and are allowing us to make progress towards an infrastructure for the exploration and understanding of petascale scientific data.

Sincerely,

Dean N. Williams

Dean N. Williams
PI, Earth System Grid Center for Enabling Technologies
Program for Climate Model Diagnosis and Intercomparison
Lawrence Livermore National Laboratory

Colorado State University
*Knowledge to Go Places*

Department of Atmospheric Science
200 W. Lake St.
Colorado State University
Fort Collins, Colorado 80523-1371

phone: 970 491-8474
fax: 970 491-8693
email: randall@atmos.colostate.edu

http://kiwi.atmos.colostate.edu/group/dave/

Monday, March 23, 2009

To Whom It May Concern:

I am writing as the Principal Investigator for SciDAC's Global Cloud Resolving Model (GCRM) project. In this role, I am responsible for developing an innovative and very high-resolution global atmospheric model.

The purpose of this letter is to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technology in their upcoming program review. In support of my SciDAC-funded science effort, the team has provided invaluable helped me and my team make progress towards petascale science by implementing a visualization and analysis strategy for the massive datasets generated by our GCRM code .

Through a productive collaborative effort with Ross Heikes (key GCRM developer at CSU) and Karen Schuhardt (PNNL), the VACET Center has had a positive impact on our (science) project. These accomplishments include:

- Development and testing of a data model for the geodesic grid and its associated variables;
- Implementation of a VisIt plugin for the model output data on the geodesic grid.

Our in-house tools are simply incapable of handling the large data volumes generated by our GCRM simulations. Thanks to the parallel capabilities and rich feature set of VisIt, we are now in a position to visualize and analyze our data and gain valuable insights.

Our group has also received support from VACET and NERSC personnel in troubleshooting and profiling large scale GCRM runs on franklin. As we ramp up our software development efforts, this effort will be extremely useful when we do 80 K+ core runs on DOE's leadership class facilities at Oak Ridge.

In summary, VACET has had a positive impact on our science effort, and is providing a valuable service to the scientific community.

Sincerely,

David A. Randall
Professor and SciDAC PI

Paul F. Fischer
Senior Scientist

Mathematics and Computer Science Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 221
Argonne, IL 60439

1-630-252-6018 phone
1-630-252-5986 fax
fischer@mcs.anl.gov

Dear SciDAC Committee Member,

This letter is to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technologies for their upcoming review. The team has provided invaluable support helping me and my collaborators make progress towards petascale science through their development and outreach for the VisIt visualization and analysis software, which we regularly use to in production mode to analyze our simulation results.

I am writing as the principal developer of the ANL fluid dynamics simulation code, Nek5000, and as an INCITE user running reactor thermal hydraulics simulations on the Blue Gene/P machine at Argonne's Advanced Leadership Computing Facility (ALCF). Our INCITE simulations regularly comprise millions of CPU hours and billions of degrees of freedom. Typical production runs involve anywhere from 8000 to 65000 processors. We use VisIt either on the ALCF analytics machine, Eureka, or on our own 128-core Linux cluster dedicated to the reactor project. Out of the many analysis options where we may have chosen to concentrate our effort, VisIt stood out as particularly attractive because of the certainty that it would scale. The *certainty* of this prospect was vital to our petascale planning as it allowed us to expend our contingency efforts in other areas.

VACET has helped us in specific ways. First, they provide infrastructure and momentum for the VisIt project, which we depend on for visualization and analysis of our petascale runs. Second, the streamline algorithms that the VACET team has developed provide a new functionality that we value, especially from an analysis perspective. Third, the networking changes VACET added to VisIt, done specifically for my collaborators, allow them to better utilize VisIt remotely. The collaboration has been so successful that we now promote VisIt as the default analysis engine for Nek5000, which is currently being used by about 30 research groups worldwide.

In short, I depend on the VisIt project for my visualization and analysis needs and VACET is helping me succeed with this project.

Sincerely,

Paul F. Fischer
Senior Scientist

**Princeton University** **Plasma Physics Laboratory**
James Forrestal Campus
P.O. Box 451, Princeton, NJ 08543

March 25, 2009

To Whom It May Concern:

The purpose of this letter is to express my strong support for the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) in their upcoming program review. The highly knowledgeable scientists of the VACET team have been developing advanced query-based visualization and analysis tools for exploring the massive data generated by our petascale kinetic applications. VACET has been working closely with our group, trying to understand exactly what we need for our analysis work in the most complex details of the physics and geometry. This is an on-going project that has already led to new incites in the physics that we are studying.

As one of the co-investigators of the SciDAC Center for Gyrokinetic Particle Simulations of Turbulent Transport in Burning Plasmas, I am responsible for the computational performance and scalability of our applications so that they can take full advantage of the petascale computing systems. Harnessing this increase in computing power ultimately leads to larger data sets to analyze from simulation containing more physics and at higher resolution. The goal is to reach a better understanding of the physics of turbulent transport in magnetic confinement fusion devices, especially in relation to ITER, the largest fusion experiment currently being built. Through a productive collaborative effort with our science team, the VACET Center has had a positive impact on our project. The initial prototype tools developed helped us better understand some of the transport phenomena linked to the nonlinear interactions between charged particles and waves in the plasma. Though the tools developed were prototypes, they already revealed a totally new way of exploring our data. We are now looking forward to the full deployment of these tools later this spring, especially now that we are about to commence simulations that will contain in excess of 20 billion particles, which will be very difficult to analyze without the VACET collaboration.

I am looking forward to continuing our collaboration as VACET has had a positive impact on our science effort, and is providing a valuable service to the scientific community.

Sincerely,

Stephane Ethier

Computational Physicist
Computational Plasma Physics Group
Princeton Plasma Physics Laboratory

Dr. John B. Bell
Center for Computational Sciences and Engineering
Computational Research Division
Lawrence Berkeley National Laboratory

E. Wes Bethel
Lawrence Berkeley National Laboratory

March 27, 2009

Dear Wes,

I'm writing this letter to express my strong support for the research and deployment activities of the Visualization and Analytics Center for Enabling Technology (VACET) in your upcoming program review. I am writing in my role as group leader for the Center for Computational Sciences and Engineering at LBNL. As part of the SciDAC program, CCSE is responsible for the development of advanced computational methods to solve large-scale scientific and engineering problems arising in Combustion as part of the APDEC ISIC and in Astrophysics as part of the Computational Astrophysics Consortium. Our primary focus is in the application of hierarchical, structured-grid finite differences for compressible, incompressible and low Mach number flows. Current applications involve low Mach number flows, such as those found in laboratory combustion experiments, and in nuclear flames in Type Ia supernovae.

The quantitative analysis of the datasets generated in our simulations has become a major challenge and our collaboration with members of the VACET team has allowed us to develop new tools based on robust topological techniques. For example we have been able for the first time to achieve a quantitative analysis of the impact of relative turbulence intensity on cellular burning structures in lean premixed hydrogen flames. This work has an impact the development of premixed burners capable of stably burning hydrogen at lean conditions. Operating at fuel-lean conditions minimizes combustion exhaust gas temperatures, which in turn reduces the formation of nitrogen-based emissions downstream of the flame. The results of this collaboration will appear in “Combustion and Flame,” the highest ranked combustion journal, in 2009.

In summary, the collaboration with the VACET team has provided a valuable contribution to our efforts to progress towards petascale science. We plan to continue and expand this collaboration and expect the development of more innovative data analysis tools that will be valuable for the broader scientific community.


Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720 | Tel: 510.486.4000

Sincerely,

Dr. John B. Bell
Senior Scientist
Lawrence Berkeley National Laboratory

**CENTER FOR TURBULENCE RESEARCH**
STANFORD UNIVERSITY
BLDG. 500, STANFORD, CALIFORNIA 94305-3035
PHONE: (650) 723-2959 • FAX: (650) 725-3525 • http://www.stanford.edu/group/ctr/

March 23, 2009

Dear SciDAC Program Reviewer,

I would like to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technologies. I am writing as a member of the SciDAC-II supported Science Application project for Simulations of Turbulent Flows with Strong Shocks and Density Variations. My colleagues and I are regular users of the VisIt visualization software and appreciate the efforts of the VACET team in helping VisIt meet our needs.

We can point to several specific ways in which VACET has helped us. First, we use the streamline capabilities that they have developed and even have images of these posted on our project website. Second, we will run on Argonne's BlueGene machine and value that VACET has gotten VisIt prepared to run on its accompanying visualization machine, Eureka. Third, we greatly value the general investments VACET has made in the VisIt project, in terms of infrastructure, general bug fixes, etc.

Summarizing, we depend on VisIt to visualize our results and we appreciate the efforts of the VACET team, which enable us to use VisIt and are also keeping VisIt prepared for the petascale future.

Sincerely,

Eric Johnsen
Post-doctoral fellow
Center for Turbulence Research
Stanford

Computing Sciences Directorate
Lawrence Berkeley National Laboratory
Berkeley, California 94720

Dr. Arie Shoshani
Mail Stop 50B-3238
Tel: (510)486-5171

March 16, 2009

To Whom It May Concern:

I'm writing this letter to express my strong support for the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) in their upcoming program review.

I am writing as the lead Principal Investigator for the SciDAC Scientific Data Management Center. In this role, I am responsible for directing the activities of a portfolio of projects aimed at solving some of the most challenging data management problems faced by the DOE science community.

Through a productive collaborative effort with our team, the VACET Center has enabled us to make substantial progress towards our objectives, namely petascale data analysis and understanding. VACET has deployed our award-winning FastBit index/query software in its production-quality, parallel-capable visual data analysis application VisIt to enable rapid exploration and knowledge discovery in some of today's largest and most complex scientific datasets. This work resulted in providing scientists with a capability to explore and visualize large datasets interactively. Since it is impossible to explore and visualize very large datasets at once, innovative visualization techniques have put to use the power of FastBit to quickly select desired subsets of the data. We have applied our joint technologies in exploring accelerator design data, climate modeling data, network traffic data to identify malicious attacks, and others. Our collaborative work with the VACET center resulted in several publications, including a field-leading publication at SC08 that included authors from the SDM Center, VACET, and the SciDAC Science Application Community Petascale Project for Accelerator Science and Simulation (ComPASS). In addition, we have shared staff, Claudio Silva from the University of Utah, who performs research, development, and deployment of data understanding workflow technology developed at the SDM center to the climate science community (the Earth Systems Grid, led by Dean Williams, LLNL).

In summary, VACET has had a positive impact on our science effort, and is providing a valuable service to the scientific community. I find the joint collaborations of the two centers very fruitful, and expect to continue and expand such collaborations in the future.

A. Shoshani

Arie Shoshani
SDM Center Director

Lawrence Livermore National Laboratory

March 18, 2009

Dear SciDAC Program Reviewer,

This letter is to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technologies. I am writing as the PI for the Interoperable Technologies for Advanced Petascale Simulations (ITAPS) center. Our center also supports VisIt development, as a vehicle to accomplish some of our own missions, and we appreciate the leverage we gain from VACET's VisIt development.

VACET has made investments in VisIt that benefit the entire VisIt community, including ITAPS. They have invested in VisIt's software engineering infrastructure. VisIt now is developed using a Subversion repository that anyone can access, whereas previously only Lawrence Livermore employees could contribute. They have stood up mailing lists for VisIt support which are archived and which are providing a greater sense of community. VACET has contributed significantly to the documentation of VisIt on the visitusers.org Wiki. Of course, they have also fixed many bugs. All of these changes make it easier for ITAPS to use VisIt and engage new users. In addition, specific tuning work done by VACET, of VisIt's subsetting mechanism ("subset inclusion lattices"),is important for our project and we benefit from this regularly.

To sum up, VACET's investment in VisIt provides momentum and leverage to a project that our center has a strong vested interest in. We appreciate their efforts and would encourage you to contact us if you have further questions.

Sincerely,

Lori Diachin

Lori Diachin
ITAPS PI





An Equal Opportunity Employer • Lawrence Livermore National Security, LLC • Operated for the US Department of Energy
P.O. Box 808 Livermore, California 94551-0808 • Telephone (925) 422-1200 • http://www.llnl.gov

David Skinner
Software Integration Group Leader
*NERSC Division, Lawrence Berkeley National Laboratory*
Principal Investigator for SciDAC Outreach

Dear SciDAC Program Reviewer,

As Principal Investigator for the Outreach Center I want to express our thanks for the work done by the VACET team in these first years of SciDAC-2. VACET has delivered solidly on two areas important to making the SciDAC program a recognized success by DOE, academia, and industry. The first regards education and outreach. The second regards technology transfer.

VACET has been and will I hope continue to be a strong contributor to the education and outreach efforts that bring SciDAC expertise to new audiences. VACET tutorials have been the main event drawing large numbers of attendees. Attendees are able to download, install, and begin using DOE software such as VisIt on their own laptops. That serves the educational goals of the tutorial, but it is important in other ways as well. It gives participants a tangible sense that DOE supported software is worthwhile, works, and can help them. They also walk away with a functioning tool they can continue to use in their research or studies.

Secondly, I also appreciate the VACET team for being patient early adopters of the collaborative development environment services provided by the SciDAC Outreach Center. These are web based code and data services targeted at helping both developers and users. They have made good use of these services and also helped us improve them by making heavy use of our source code controls systems, reporting bugs and making service requests. By doing so they have helped move the expectations for DOE HPC software in a direction that enables technology transfer from SciDAC to the rest of the world.

Please feel free to contact me if there are any questions or concerns.

David Skinner
Software Group Leader, LBNL
Principal Investigator, SciDAC Outreach


Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720 | Tel: 510.486.4000

Lawrence Livermore National Laboratory

March 3, 2009

Dear SciDAC Reviewer,

This letter is to express support for the Visualization and Analytics Center for Enabling Technologies (VACET) for their role in the development of the visualization and analysis tool VisIt. VisIt's primary role is to process the extremely large data sets made by simulation codes on massively parallel machines, so VACET's contributions to improving VisIt to work better on large data sets is much appreciated.

VACET's major development efforts have been great successes and have contributed greatly to the VisIt project as a whole. The parallel streamline algorithms they developed are cutting edge and filled what was previously a weak spot in VisIt. These algorithms have gone on to be used by customers outside the Office of Science and we often use this work as an example for the benefits of partnering on VisIt development. The work using accelerated queries with the FastBit library adds on to our existing investments in processing large data and demonstrates how well aligned VACET development is with VisIt's identity. Of course, there have been many, many other enhancements, bug fixes, porting fixes, and infrastructural work that have benefited users of VisIt all over the world. (VisIt has been downloaded more than 100,000 times.)

Another major contribution by VACET has been to make VisIt appear more like an open source project. Previous to VACET, Lawrence Livermore computer scientists served as the primary VisIt developers with the NNSA comprising the majority of the tool's customer base. When VACET was funded, they immediately began changing our project's culture to be more "friendly" to new developers. We now have a publicly available repository, Wikis, open mailing lists that are archived and other trappings of open source projects. These changes, the majority of which were done directly by VACET developers, have in turn led to more partnerships, including a partnership with the Office of Nuclear Energy, universities, and foreign laboratories. Without VACET, these developments were much less likely.

In summary, let me again state that VACET has provided a significant boost to our project and they are enabling new features for many users. In addition to being the general VisIt project lead, I also lead the NNSA VisIt efforts and I believe that VACET development provides significant leverage to the NNSA efforts.

Best regards,

Eric Brugger

Eric Brugger
VisIt project leader

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

One Bethel Valley Road
P.O. Box 2008, MS-6008
Oak Ridge, TN 37831-6016
(865) 241-3748

3 March 2008

Dear SciDAC reviewer:

I am writing in my capacity as the Director of Science for Oak Ridge National Laboratory's National Center for Computational Sciences (NCCS) to voice my strong support for the VACET Visualization and Analytics Center for Enabling Technologies. As a steward of high performance scientific computing at one of our nation's premier computing centers, I value the role of scientific visualization and analysis in the exploration of data that our users generate. The tools that VACET members produce, as well as the support that they give to SciDAC users on ORNL's systems, greatly complement the capabilities that NCCS provides.

For the last 2.5 years, VACET members have deployed the VisIt visualization tool on our systems. They have maintained it for SciDAC users, providing support and producing images that we regularly use in our publications and public relations materials. In addition, the parallel R capabilities that have been deployed on our systems have enabled parallel statistical analysis that would not have been possible elsewhere, especially for climate science exploration. While we have an embedded visualization team within our center, their capabilities are greatly strengthened by the collaborations that SciDAC and VACET provide.

We at NCCS are very pleased with the services that VACET provides for our users and we look forward to future partnerships with them.

Sincerely,

Douglas B. Kothe
Director of Science
Leadership Computing Facility
National Center for Computational Sciences
Oak Ridge National Laboratory
*kothe@ornl.gov*

BERKELEY LAB

NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

March 23, 2009

To Whom It May Concern:

I am writing in my capacity as the Director of the National Energy Research Scientific Computing Center (NERSC) in support of the VACET Visualization and Analytics Center for Enabling Technologies. As DOE's flagship computing center, NERSC provides high end computing, storage and data analysis facilities to all of the programs within the Office of Science. Scientific analysis and visualization of both measured and simulated data provide a essential component of computational science, and the tools that VACET members produce are critical to the services provided by NERSC. The VACET team works closely with the NERSC Analytics groups, providing access to the latest tools available from the SciDAC program and direct feedback between the research and development of visualization tools and the use of those tools in production applications. Production-quality, parallel capable software from VACET, like the VisIt application, is installed and in use at NERSC today. Additionally, VACET has worked and continues to work with several key science teams and the NERSC Analytics team to apply these tools on NERSC systems to address challenging scientific knowledge discovery problems.

I view the VACET collaboration as an important part of our overall data analytics strategy at NERSC, and we are very pleased with the interactions between NERSC users and the VACET team.

Sincerely,

Katherine Yelick

Katherine Yelick
Professor of EECS, University of California at Berkeley
NERSC Director, Lawrence Berkeley National Laboratory



LAWRENCE BERKELEY NATIONAL LABORATORY
ONE CYCLOTRON ROAD MS 50B 4230|BERKELEY, CALIFORNIA 94720|TEL: 510.495.2431|510.486.4300

William E. Allcock
Manager, Advanced Integration Group
Acting Director of Operations

Leadership Computing Facility
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 360
Argonne, IL 60439-4844

1-630-252-7573 office
1-630-252-2828 fax
allcock@alcf.anl.gov

Dear SciDAC Reviewer:

I'm writing this review letter to express strong support for the SciDAC Visualization and Analytics Center for Enabling Technology (VACET) in their upcoming program review.

I am writing as the Manager of the Advanced Integration Group and Acting Director of Operations for the Argonne Leadership Computing Facility (ALCF). In this role, one of my responsibilities is to ensure that the visualization and analysis needs are met for the INCITE simulations run on ALCF machines.

We greatly appreciate the support we have received from VACET in getting VisIt up and running at Argonne. We have had many INCITE users request VisIt for their visualization and analysis needs and VACET has been crucial to meeting these needs. They have ported, installed, and maintained VisIt on Eureka, our dedicated machine for analyzing the results coming off our BG/P machine. They have modified VisIt's communication protocols to enable client-server communication specifically for our Eureka machine, which enables our many remote users to use VisIt effectively. They have even assisted in porting the Silo I/O library, because of the Silo expertise on their team. Finally, they have been willing to consult with ALCF team members on how to effectively use VisIt. All of these efforts have been important to the success of our mission.

Please feel free to contact me should you need any further information.

Best Regards,

William E. Allcock
Manager, Advanced Integration Group
Acting Director of Operations

WE:dmn